second edition

lectures on quantum mechanics

basic matters

berthold-georg englert

lectures on quantum mechanics

basic matters

second edition

Other Lecture Notes by the Author

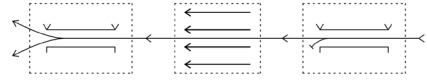
Lectures on Classical Electrodynamics
ISBN: 978-981-4596-92-3
ISBN: 978-981-4596-93-0 (pbk)

Lectures on Classical Mechanics
ISBN: 978-981-4678-44-5
ISBN: 978-981-4678-45-2 (pbk)

Lectures on Statistical Mechanics
ISBN: 978-981-12-2457-7
ISBN: 978-981-122-554-3 (pbk)

lectures on quantum mechanics

basic matters



second edition

Berthold-Georg Englert
Beijing Institute of Technology, China
National University of Singapore, Singapore

World Scientific

NEW JERSEY • LONDON • SINGAPORE • BEIJING • SHANGHAI • HONG KONG • TAIPEI • CHENNAI • TOKYO

Published by

World Scientific Publishing Co. Pte. Ltd.

5 Toh Tuck Link, Singapore 596224

USA office: 27 Warren Street, Suite 401-402, Hackensack, NJ 07601

UK office: 57 Shelton Street, Covent Garden, London WC2H 9HE

Library of Congress Cataloging-in-Publication Data
Names: Englert, Berthold-Georg, 1953– author.
Title: Lectures on quantum mechanics / Berthold-Georg Englert.
Description: Second edition, corrected and enlarged. | Hackensack : World Scientific Publishing Co. Pte. Ltd., 2024. | Includes bibliographical references and index. |
Contents: Basic matters -- Simple systems -- Perturbed evolution.
Identifiers: LCCN 2023040959 (print) | LCCN 2023040960 (ebook) |
ISBN 9789811284724 (v. 1 ; hardcover) | ISBN 9789811284984 (v. 1 ; paperback) |
ISBN 9789811284755 (v. 2 ; hardcover) | ISBN 9789811284991 (v. 2 ; paperback) |
ISBN 9789811284786 (v. 3 ; hardcover) | ISBN 9789811285004 (v. 3 ; paperback) |
ISBN 9789811284731 (v. 1 ; ebook) | ISBN 9789811284762 (v. 2 ; ebook) |
ISBN 9789811284793 (v. 3 ; ebook)
Subjects: LCSH: Quantum theory. | Physics.
Classification: LCC QC174.125 .E54 2023 (print) | LCC QC174.125 (ebook) |
DDC 530.12--dc23/eng/20231011
LC record available at https://lccn.loc.gov/2023040959
LC ebook record available at https://lccn.loc.gov/2023040960

British Library Cataloguing-in-Publication Data
A catalogue record for this book is available from the British Library.

Copyright © 2024 by World Scientific Publishing Co. Pte. Ltd.

All rights reserved. This book, or parts thereof, may not be reproduced in any form or by any means, electronic or mechanical, including photocopying, recording or any information storage and retrieval system now known or to be invented, without written permission from the publisher.

For photocopying of material in this volume, please pay a copying fee through the Copyright Clearance Center, Inc., 222 Rosewood Drive, Danvers, MA 01923, USA. In this case permission to photocopy is not required from the publisher.

For any available supplementary material, please visit
https://www.worldscientific.com/worldscibooks/10.1142/13636#t=suppl

Printed in Singapore

To my teachers, colleagues, and students

This page intentionally left blank

Preface

This book on the *Basic Matters* of quantum mechanics grew out of a set of lecture notes for a second-year undergraduate course at the National University of Singapore (NUS). It is a first introduction that does not assume any prior knowledge of the subject. The presentation is rather detailed and does not skip intermediate steps that — as experience shows — are not so obvious for the learning student.

Starting from the simplest quantum phenomenon, the Stern–Gerlach experiment with its choice between two discrete outcomes, and ending with the standard examples of one-dimensional continuous systems, the physical concepts and notions as well as the mathematical formalism of quantum mechanics are developed in successive small steps, with scores of exercises along the way. The presentation is "modern," a dangerous word, in the sense that the natural language of the trade — Dirac's kets and bras and all that — is introduced early, and the temporal evolution is dealt with in a picture-free manner, with Schrödinger's and Heisenberg's equations of motion side by side and on equal footing.

Two companion books on *Simple Systems* and *Perturbed Evolution* cover the material of the subsequent courses at NUS for third- and fourth-year students, respectively. The three books are, however, not strictly sequential but rather independent of each other and largely self-contained. In fact, there is quite some overlap and a considerable amount of repeated material. While the repetitions send a useful message to the self-studying reader about what is more important and what is less, one could do without them and teach most of *Basic Matters*, *Simple Systems*, and *Perturbed Evolution* in a coherent two-semester course on quantum mechanics.

All three books owe their existence to the outstanding teachers, colleagues, and students from whom I learned so much. I dedicate these lectures to them.

I am grateful for the encouragement of Professors Choo Hiap Oh and Kok Khoo Phua who initiated this project. The professional help by the staff of World Scientific Publishing Co. was crucial for the completion; I acknowledge the invaluable support of Miss Ying Oi Chiew and Miss Lai Fun Kwong with particular gratitude. But nothing would have come about, were it not for the initiative and devotion of Miss Jia Li Goh who turned the original handwritten notes into electronic files that I could then edit.

I wish to thank my dear wife Ola for her continuing understanding and patience by which she is giving me the peace of mind that is the source of all achievements.

Singapore, March 2006 *BG Englert*

Note on the second edition

The feedback received from students and colleagues, together with my own critical take on the three companion books on quantum mechanics, suggested rather strongly that the books would benefit from a revision. This task has now been completed.

Many readers have contributed entries to the list of errata. I wish to thank all contributors sincerely and extend special thanks to Miss Hong Zhenxi and Professor Lim Hock.

In addition to correcting the errors, I tied up some loose ends and brought the three books in line with the later volumes in the "Lectures on ..." series. There is now a glossary, and the exercises, which were interspersed throughout the text, are collected after the main chapters and supplemented by hints.

The team led by Miss Nur Syarfeena Binte Mohd Fauzi at World Scientific Publishing Co. contributed greatly to getting the three books into shape. I thank them very much for their efforts.

Beijing and Singapore, November 2023 *BG Englert*

Contents

Preface **vii**

Glossary **xiii**
Miscellanea xiii
Latin alphabet xiv
Greek alphabet and Greek-Latin combinations xv

1. A Brutal Fact of Life **1**
1.1 Causality and determinism 1
1.2 Bell's inequality: No hidden determinism 4
1.3 Remarks on terminology 7

2. Kinematics: How Quantum Systems are Described **9**
2.1 Stern–Gerlach experiment 9
2.2 Successive Stern–Gerlach measurements 12
2.3 Order matters 14
2.4 Mathematization 15
2.5 Probabilities and probability amplitudes 20
2.6 Quantum Zeno effect 29
2.7 Kets and bras 31
2.8 Brackets, bra-kets, and ket-bras 34
2.9 Pauli operators, Pauli matrices 38
2.10 Functions of Pauli operators 40
2.11 Eigenvalues, eigenkets, and eigenbras 42
2.12 Wave–particle duality 46
2.13 Expectation value 47
2.14 Trace 49

2.15 Statistical operator, Born rule 51
2.16 Mixtures and blends 55
2.17 Nonselective measurement 55
2.18 Entangled atom pairs 57
2.19 State reduction, conditional probabilities 64
2.20 Measurement outcomes do not pre-exist 66
2.21 Measurements with more than two outcomes 68
2.22 Unitary operators 73
2.23 Hermitian operators 76
2.24 Hilbert spaces for kets and bras 77

3. Dynamics: How Quantum Systems Evolve 79
3.1 Schrödinger equation 79
3.2 Heisenberg equation 83
3.3 Equivalent Hamilton operators 85
3.4 Von Neumann equation 86
3.5 Example: Larmor precession 87
3.6 Time-dependent probability amplitudes 90
3.7 Schrödinger equation for probability amplitudes 91
3.8 Time-independent Schrödinger equation 95
3.9 Example: Two magnetic silver atoms 97

4. Motion Along the x Axis 105
4.1 Kets, bras, and wave functions 105
4.2 Position operator 110
4.3 Momentum operator 112
4.4 Heisenberg's commutation relation 114
4.5 Position–momentum transformation function 115
4.6 Expectation values 117
4.7 Uncertainty relation 120
4.8 State of minimum uncertainty 123
4.9 Time dependence 126
4.10 Excursion into classical mechanics 127
4.11 Hamilton operator, Schrödinger equation 131
4.12 Time transformation function 133

5. Elementary Examples **135**
5.1 Force-free motion . . . 135
5.1.1 Time-transformation functions . . . 135
5.1.2 Spreading of the wave function . . . 137
5.1.3 Long-time and short-time behavior . . . 142
5.1.4 Interlude: General position-dependent force . . . 148
5.1.5 Energy eigenstates . . . 150
5.2 Constant force . . . 153
5.2.1 Energy eigenstates . . . 153
5.2.2 Limit of no force . . . 155
5.3 Harmonic oscillator . . . 158
5.3.1 Energy eigenstates: Power-series method . . . 158
5.3.2 Energy eigenstates: Ladder-operator approach . . . 163
5.3.3 Hermite polynomials . . . 168
5.3.4 Infinite matrices . . . 170
5.3.5 Position and momentum spreads . . . 172
5.4 Delta potential . . . 173
5.4.1 Bound state . . . 173
5.4.2 Scattering states . . . 178
5.5 Square-well potential . . . 182
5.5.1 Bound states . . . 182
5.5.2 Delta potential as a limit . . . 186
5.5.3 Scattering states and tunneling . . . 187
5.6 Stern–Gerlach experiment revisited . . . 191

Exercises with Hints **195**
Exercises for Chapters 1–5 . . . 195
Hints . . . 209

Index **217**

This page intentionally left blank

Glossary

Here is a list of the symbols used in the text; the numbers in square brackets indicate the pages of first occurrence or of other significance.

Miscellanea

0	null symbol: number 0, or null column, or null matrix, or null ket, or null bra, or null operator, *et cetera*
1	unit symbol: number 1, or unit matrix, or identity operator, *et cetera*
$A \mathrel{\widehat{=}} B$	read "A represents B" or "A is represented by B"
$\mathrm{Max}\{\ \}, \mathrm{Min}\{\ \}$	maximum, minimum of a set of real numbers
$(\ ,\ )$	inner product [77]
a^*, $\lvert a \rvert$	complex conjugate of a, absolute value of a
$\mathrm{Re}(a), \mathrm{Im}(a)$	real, imaginary part of a: $a = \mathrm{Re}(a) + \mathrm{i}\,\mathrm{Im}(a)$
$a = \lvert \boldsymbol{a} \rvert$	length of vector $\boldsymbol{a}$
$\boldsymbol{a} \cdot \boldsymbol{b}, \boldsymbol{a} \times \boldsymbol{b}$	scalar, vector product of vectors $\boldsymbol{a}$ and $\boldsymbol{b}$
$A^\dagger$	adjoint of A [25]
$\det(A), \mathrm{tr}(A)$	determinant [43], trace of A [49]
$\lvert\ \rangle, \langle\ \rvert$; $\lvert 1 \rangle, \langle a \rvert$	generic ket, bra; labeled ket, bra [33]
$\lvert \ldots, t \rangle$, $\langle \ldots, t \rvert$	ket, bra at time t [81]
$\langle\ \rvert\ \rangle, \lvert\ \rangle\langle\ \rvert$	bra-ket, ket-bra [35]
$\langle \ldots, t_1 \rvert \ldots, t_2 \rangle$	time transformation function [134]
$\langle A \rangle$	mean value, expectation value of A [48]
$[A, B]$	commutator of A and B [83]
$\uparrow_z, \downarrow_z$	spin-up, spin-down in the z direction [33]
$x!$	factorial of x [112]
$f^2(x)$, $f^{-1}(x)$	square, inverse of the function $x \mapsto f(x)$: $f^2(x) = f\big(f(x)\big)$, $f\big(f^{-1}(x)\big) = x$, $f^{-1}\big(f(x)\big) = x$

$f(x)^2$, $f(x)^{-1}$	square, reciprocal of the function value: $f(x)^2 = \big(f(x)\big)^2$, $f(x)^{-1} = 1/f(x)$
$\mathrm{d}t$, δt	differential, variation of t
$\dfrac{\mathrm{d}}{\mathrm{d}t}$, $\dfrac{\partial}{\partial t}$	total, parametric time derivative
$\boldsymbol{\nabla}$	gradient vector differential operator
$\otimes$	tensor product: $\|a\rangle \otimes \|b\rangle = \|a, b\rangle$ [59]

Latin alphabet

a	width of the square-well potential [182]
a, b	detector settings [4]
a_{jk}, b_{jk}	matrix elements [15]
$a_k^{(n)}$	power series coefficients [160]
$A, A^\dagger$	harmonic-oscillator ladder operators [163]
A, B	matrices [15]
$\boldsymbol{B}$	magnetic field [10]
$B_0(t)$, $b(t)$	on-axis field strength, near-axis gradient [191]
$C(a, b)$	Bell correlation [5]
$\cos, \sin, \ldots$	trigonometric functions
$\cosh, \sinh, \ldots$	hyperbolic functions
e; $\mathrm{e}^x = \exp(x)$	Euler's number, $\mathrm{e} = 2.71828\ldots$; exponential function
$\boldsymbol{e}, \boldsymbol{n}$	unit vectors
$\boldsymbol{e}_x$, $\boldsymbol{e}_y$, $\boldsymbol{e}_z$	unit vectors for the x, y, z directions [32]
E_{magn}; E_k	magnetic energy [10]; kth eigenenergy [95]
F	flipper [17]
$\boldsymbol{F}, F$	force [11,128]
G_a, G_b	generators [128]
$h = 2\pi\hbar$	Planck's constant, $\hbar = 1.05457 \times 10^{-34}\,\mathrm{J\,s} = 0.658212\,\mathrm{eV\,fs}$ [82]
$H, \mathcal{H}$	Hamilton operator [81], matrix representing H [92]
$\mathrm{H}_n(\ )$	nth Hermite polynomial [162]
i	imaginary unit, $\mathrm{i}^2 = -1$
L	coherence length [206]; Lagrange function [127]
M	mass [128]
$O\big(t^4\big)$	terms of oder t^4 or smaller [144]
prob(E)	probability for event E [27]
$p; \overline{p}$	momentum [115]; at the stationary phase [156]
$P, P(t)$	momentum operator [113], with time dependence [126]

r	distance between atoms [97]
q	distance in oscillator units [159]
r, t	reflected, transmitted amplitude [188]
R,T	single-photon counters [3]
$R_y(\ )$	rotation matrix for the y axis [20]
$\boldsymbol{s}$	Bloch vector [54]
$\mathrm{sgn}(x)$	sign of x
S_+, S_-	+selector, −selector [16]
t; t_0, t_1	time [81]; time before, after [80]
$U_{ab}, U_t(\tau), \ldots$	unitary operators [73]
V_0	depth of the square-well potential [182]
$V(x)$	potential energy [128]
w_k	weights in blending ρ [121]
$W_{ab}; W(p)$	action [127]; action variable [155]
$x, y, z; \dot{x}$	cartesian coordinates [9]; velocity [127]
X, Z	position operator for the x, z direction [111]
$X(t), Z(t)$	time-dependent position operators [126]

Greek alphabet and Greek-Latin combinations

α	phase shift [178]
$\alpha, \beta, \ldots; \alpha_k$	probability amplitudes [16]; kth amplitude [96]
δ_{jk}	Kronecker's delta symbol [69]
$\delta(x - x')$	Dirac's delta function [106]
$\delta(x - x'; \epsilon)$	model for Dirac's delta function [108]
$\delta A, (\delta A)^2$	spread, variance of A [120]
$\Delta p(t)$, $\Delta z(t)$	transferred momentum, associated displacement [193]
$\epsilon(t)$	complex spreading factor [140]
$\eta(x)$	Heaviside's step function [208]
θ, ϑ	energy parameters (square well) [183]
ϑ	rotation angle [29]
κ	reciprocal length [173]
κ, k	energy parameters (square well) [182]
λ	hidden variable [5]; eigenvalue candidate [22]
λ	de Broglie wavelength [117]
λ_+, λ_-	measurement results [47]
$A_\lambda(a), B_\lambda(b)$	measurement results [6]
$w(\lambda)$	hidden-variable weight density [5]
Λ	measurement symbol [47]

$\mu_{\rm B}; \boldsymbol{\mu}$	Bohr magneton [202]; magnetic dipole moment vector [10]
π	Archimedes's constant, $\pi = 3.14159\ldots$
$\rho, \rho^{(\rm pair)}, \rho_{\rm T}^{(1)}, \ldots$	statistical operators [52]
$\sigma_x, \sigma_y, \sigma_z; \boldsymbol{\sigma}$	Pauli operators for the x, y, z direction [39]; Pauli vector operator [40]
$\boldsymbol{\sigma}^{(1)}, \boldsymbol{\sigma}^{(2)}$	Tom's, Jerry's Pauli vector operator [57]
τ	small time increment [81], path parameter [128]
$t(\tau), x(\tau)$	path variables [128]
$\boldsymbol{\tau}$	torque [10]
ϕ	rotation angle [20]; complex phase [28]
$\phi_k, \phi_k^\dagger$	kth eigencolumn, its adjoint [96]
φ	complex phase [28]
φ, ϑ	azimuth, polar angle [44]
Φ	hermitian phase operator [76]
$\chi_n(q)$	wave function factor [160]
$\psi(t)$	time-dependent column of probability amplitudes [94]
$\psi(x), \psi(p)$	position [105], momentum wave function [118]
$\omega, \Omega; \boldsymbol{\omega}, \boldsymbol{\Omega}$	angular frequencies; vectors [85,87]
$\omega; \omega_{kl}$	oscillator frequency [158]; transition frequency [97]

Chapter 1

A Brutal Fact of Life

1.1 Causality and determinism

Before their first encounter with the quantum phenomena that govern the realm of atomic physics and sub-atomic physics, students receive a training in classical physics, where Newton's* mechanics of massive bodies and Maxwell's† electromagnetism — the physical theory of the electromagnetic field and its relation to the electric charges — give convincingly accurate accounts of the observed phenomena. Indeed, almost all experiences of physical phenomena that we are conscious of without the help of refined instruments fit perfectly into the conceptual and technical framework of these classical theories. It is instructive to recall two characteristic features that are equally possessed by Newton's mechanics and Maxwell's electromagnetism: *Causality* and *Determinism*.

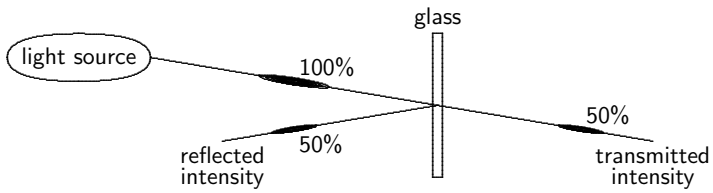
Causality is inference in time: Once you know the state of affairs — physicists prefer to speak more precisely of the "state of the system" — you can predict the state of affairs at any later time, and often also retrodict the state of affairs at earlier times. Having determined the relative positions of the sun, earth, and moon and their relative velocities, we can calculate highly precisely when the next lunar eclipse will happen (extreme precision on short time scales and satisfactory precision for long time scales also require good knowledge of the positions and velocities of the other planets and their satellites, but that is a side issue here) or when the last one occurred. Quite similarly, present knowledge of the strength and direction of the electric and magnetic fields together with knowledge about the motion of the electric charges enable us to calculate reliably the electromagnetic field configuration in the future or the past.

*Isaac Newton (1643–1727) †James Clerk Maxwell (1831–1879)

Causality, as we shall see, is also a property of quantal evolution: Given the state of the system now, we can infer the state of the system later (but, typically, not earlier). Such as there are Newton's equation of motion in mechanics and Maxwell's set of equations for the electromagnetic field, there are also equations of motion in quantum mechanics: Schrödinger's* equation, which is more in the spirit of Maxwell's equations, and Heisenberg's† equation, which is more in Newton's tradition.

We say that the classical theories are *deterministic* because the state of the system uniquely determines all phenomena. When the positions and velocities of all objects are known in Newton's mechanics, the results of all possible measurements are predictable, there is no room for any uncertainty in principle. Likewise, once the electromagnetic field is completely specified and the positions and velocities of all charges are known in Maxwell's electromagnetic theory, all possible electromagnetic phenomena are fully predictable.

Let us look at a somewhat familiar situation that illustrates this point and will enable us to establish the difference in situation that we encounter in quantum physics. You have all seen reflections of yourself in the glass of a shopping window while at the same time having a good view of the merchandise for sale. This is a result of the property of the glass sheet that it partly transmits light and partly reflects it. In a laboratory version, we could have 50% probability each for transmission and reflection:



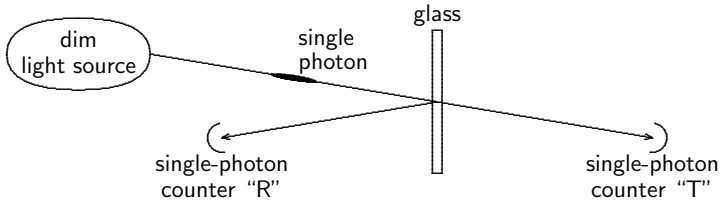
(1.1.1)

A light source emits pulses of light, which are split in two by such a half-transparent mirror, half of the intensity being transmitted and the other half reflected. Given the properties of pulses emitted by the source and the material properties of the glass, we can predict completely how much of the intensity is reflected, how much is transmitted, how the pulse shape is changed, and so forth — all these are implications of Maxwell's equations.

But, we know that there is a different class of phenomena that reveal a certain graininess of light: the pulses consist of individual lumps of energy — "light quanta" or "photons." (We are a bit sloppy with the terminology

*Erwin Schrödinger (1887–1961) †Werner Heisenberg (1901–1976)

here; at a more refined level, photons and light quanta are not the same, but that is irrelevant presently.) We become aware of the photons if we dim the light source by so much that there is only a single photon per pulse. We also register the reflected and transmitted light by single-photon counters:



(1.1.2)

What will be the fate of the next photon to arrive? Since it cannot split in two, either the photon is transmitted as a whole or it is reflected as a whole so that eventually *one* of the counters will register the photon. A single photon, so to say, makes one detector click: either we register a click of detector T or of detector R, but not of both.

What is important here is that we *cannot predict* which detector will click for the next photon, all we know is the history of the clicks of the photons that have already arrived. Perhaps a sequence such as

$$\mathsf{R\ R\ T\ R\ T\ T\ T\ R\ T\ R} \tag{1.1.3}$$

was the case for the last ten photons. In a long sequence, reporting the detector clicks of very many photons, there will be about the same number of T clicks and R clicks because it remains true that half of the intensity is reflected and half transmitted. On the single-photon level, this becomes a *probabilistic* fact: Each photon has a 50% chance of being reflected and an equal chance of being transmitted. And this is *all* we can say about the future fate of a photon approaching the glass sheet.

So, when repeating the experiment with another set of ten photons, we do not reproduce the above sequence of detector clicks, but rather get another one, perhaps

$$\mathsf{R\ T\ T\ R\ R\ R\ R\ T\ R\ T}. \tag{1.1.4}$$

And a third set of ten would give yet another sequence, all 2^{10} possible sequences occurring with the same frequency if we repeat the experiment very often.

Thus, although we know exactly all the properties of the incoming photon, we cannot predict which detector will click. We can only make statis-

tical predictions that answer questions such as the following: "How likely are four Ts and six Rs in the next sequence of ten?"

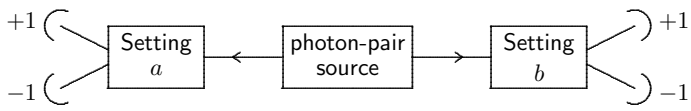
What we face here, in a simple but typical situation, is the *lack of determinism* of quantum phenomena. Complete knowledge of the state of affairs does not enable us to predict the outcomes of all measurements that could be performed on the system. In other words, the state does not determine the phenomena. There is a fundamental element of chance: The laws of nature that connect our knowledge about the state of the system with the observed phenomena are *probabilistic*, not deterministic.

1.2 Bell's inequality: No hidden determinism

Now, that raises the question about the origin of this probabilistic nature: Does the lack of determinism result from a lack of knowledge? Or, put differently, could we know more than we do and then have determinism reinstalled? The answer is *No*. Even if we know everything that can possibly be known about the photon, we cannot predict its fate.

It is not simple to make this point for the example discussed above with single photons incident on a half-transparent mirror. In fact, one can construct contrived formalisms in which the photons are equipped with internal clockworks of some sort that determine in a hidden fashion where each photon will go. But in more complicated situations, even the most ingenious deterministic mechanism cannot reproduce the observed facts in all respects. The following argument is a variant of the one given by Bell* in the 1960s.

Consider the more general scenario in which a photon-pair source always emits two photons, one going to the left and the other going to the right:



(1.2.1)

Each photon is detected by one of two detectors eventually — with measurement results +1 or −1 — and the devices allow for a number of parameter settings. We denote by symbol a the collection of parameters on the left, and by b those on the right. Details do not matter; we just need that different settings are possible, that there is a choice between different mea-

*John Stewart Bell (1928–1990)

surements on both sides. The only restriction we insist upon is that there are only two possible outcomes for each setting, the abstract generalization of "transmission" and "reflection" in the single-photon plus glass sheet example above.

For any given setting, the experimental data are of this kind:

photon pair no.	1	2	3	4	5	6	7	8	...
on the left	+1	+1	−1	−1	+1	−1	+1	+1	...
on the right	+1	−1	−1	+1	−1	−1	−1	+1	...
product	+1	−1	+1	−1	−1	+1	−1	+1	...

(1.2.2)

The products in the last row distinguish the pairs with the same outcomes on the left and the right (product = +1) from those with opposite outcomes (product = −1). We use these products to define the *Bell correlation* $C(a, b)$ for the chosen setting specified by parameters a, b,

$$C(a,b) = \frac{(\text{number of } +1 \text{ pairs}) - (\text{number of } -1 \text{ pairs})}{\text{total number of observed pairs}}\,. \tag{1.2.3}$$

Clearly, we have $C(a, b) = +1$ if $+1$ on one side is always matched with a $+1$ on the other and -1 with -1, and we have $C(a, b) = -1$ if $+1$ is always paired with -1 and -1 with $+1$. In all other cases, the value of $C(a, b)$ is between these extrema, so that

$$-1 \le C(a,b) \le +1 \tag{1.2.4}$$

for any setting a, b.

Following Bell, let us now fantasize about a (hidden) mechanism that determines the outcome on each side. We conceive each pair as being characterizable by a set of parameters collectively called λ and that the source realizes various λ with different relative frequencies. Thus, there is a positive weight function $w(\lambda)$, such that $\mathrm{d}\lambda\, w(\lambda)$ is the probability of having a λ value within a $\mathrm{d}\lambda$ volume around λ. These probabilities must be positive numbers that sum up to unity,

$$w(\lambda) \ge 0\,, \qquad \int \mathrm{d}\lambda\, w(\lambda) = 1\,. \tag{1.2.5}$$

We need not be more specific because further details are irrelevant to the argument — which is, of course, the beauty of it.

We denote by $A_\lambda(a)$ the measurement result on the left for setting a when the hidden control parameter has value λ and by $B_\lambda(b)$ the corresponding measurement result on the right. Since all measurement results

are either $+1$ or -1, we have

$$A_\lambda(a) = \pm 1\,, \qquad B_\lambda(b) = \pm 1 \qquad \text{for all} \quad a, b, \lambda \tag{1.2.6}$$

and also

$$A_\lambda(a) B_\lambda(b) = \pm 1 \qquad \text{for all} \quad a, b, \lambda\,. \tag{1.2.7}$$

This is then the product to be entered in the table (1.2.2) for the pair that leaves the source with value λ and encounters the settings a and b. Upon summing over all pairs, we get

$$C(a,b) = \int \mathrm{d}\lambda\, w(\lambda) A_\lambda(a) B_\lambda(b) \tag{1.2.8}$$

for the Bell correlation, and all the rest follows from this expression.

Before proceeding, however, let us note that an important assumption has entered: We take for granted that the measurement result on the left does not depend on the setting of the apparatus on the right, and *vice versa*. This is an expression of *locality* as we naturally accept it as a consequence of Einstein's* observation that spatially well-separated events cannot be connected by any causal links if they are simultaneous in one reference frame. Put differently, if the settings a and b are decided very late, just before the measurements actually take place, any influence of the setting on one side upon the outcome on the other side would be inconsistent with Einsteinian causality. With this justification, there is no need to consider the more general possibility of having $A_\lambda(a,b)$ on the left and $B_\lambda(a,b)$ on the right. Such a b dependence of A_λ and an a dependence of B_λ are physically unacceptable, but of course, it remains a mathematical possibility that cannot be excluded on purely logical grounds.

All together, we now consider two settings on the left, a and a', and two on the right, b and b'; then, there are four Bell correlations. We subtract the fourth from the sum of the other three,

$$\begin{aligned} &C(a,b) + C(a,b') + C(a',b) - C(a',b') \\ &= \int \mathrm{d}\lambda\, w(\lambda) \Big[A_\lambda(a) \underbrace{\big(B_\lambda(b) + B_\lambda(b')\big)}_{=0 \text{ or } \pm 2} + A_\lambda(a') \underbrace{\big(B_\lambda(b) - B_\lambda(b')\big)}_{=0 \text{ or } \pm 2} \Big]\,, \end{aligned} \tag{1.2.9}$$

where, for each value of λ, one of the $B_\lambda(b) \pm B_\lambda(b')$ terms equals 0 and the other term is ± 2 as follows from (1.2.6). Therefore, $\big[\cdots\big]$ in (1.2.9) is ± 2

*Albert Einstein (1879–1955)

for every λ, and since the integral is the average value, we conclude that

$$\Big|C(a,b)+C(a,b')+C(a',b)-C(a',b')\Big| \leq 2\,. \tag{1.2.10}$$

This is (a variant of) the so-called *Bell inequality*. Given the very simple argument and the seemingly self-evident assumptions entering at various stages, one should confidently expect that it is generally obeyed by the correlations observed in any experiment of the kind depicted in (1.2.1). Anything else would defy common sense, would it not? But the fact is that rather strong violations *are* observed in real-life experiments in which the left-hand side substantially exceeds 2, getting very close indeed to $2\sqrt{2}$, the maximal value allowed for Bell correlations in quantum mechanics (see Section 2.18).

Since we cannot possibly give up our convictions about locality, and thus about Einsteinian causality, the logical conclusion must be that there just is no such hidden deterministic mechanism [characterized by $w(\lambda)$ as well as $A_\lambda(a)$ and $B_\lambda(b)$]. We repeat:

$$\begin{array}{l}\text{There is no mechanism that decides the}\\ \text{outcome of a quantum measurement.}\end{array} \tag{1.2.11}$$

What is true for such correlated pairs of photons is, by inference, also true for individual photons. There is no mechanism that decides whether the photon is transmitted or reflected by the glass sheet; it is rather a *truly probabilistic* phenomenon.

This is a brutal fact of life. In a very profound sense, quantum mechanics is about learning to live with it.

1.3 Remarks on terminology

We noted the fundamental lack of determinism at the level of quantum phenomena and the consequent inability to predict the outcome of all experiments that could be performed. It may be worth emphasizing that this lack of predictive power is of a very different kind than, say, the impossibility of forecasting next year's weather.

The latter is a manifestation of the chaotic features of the underlying dynamics, frequently referred to as *deterministic chaos*. In this context, "deterministic" means that the equations of motion are differential equations that have a unique solution for given initial values — the property that we called "causal" above. There is a clash of terminology here if one wishes to diagnose one.

The deterministic chaos comes about because the solutions of the equations of motion depend extremely sensitively on the initial values, which in turn are never known with utter precision. This sensitivity is a generic feature of nonlinear equations and not restricted to classical phenomena. Heisenberg's equations of motion of an interacting quantum system are just as nonlinear as Newton's equations for the corresponding classical system if there is one.

In classical systems that exhibit deterministic chaos, our inability to make reliable predictions concerns phenomena that are sufficiently far away in the future — a weather forecast for the next three minutes is not such a challenge. In the realm of quantum physics, however, the lack of determinism is independent of the time elapsed since the initial conditions were established. Even perfect knowledge of the state of affairs immediately before a measurement is taken does not enable us to predict the outcome; at best, we can make a probabilistic prediction, a statistical prediction.

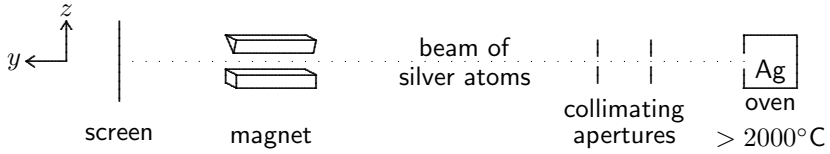
Of course, matters tend to be worse whenever one extrapolates from the present situation, which is perhaps known with satisfactory precision, to future situations. The knowledge may not be accurate enough for a long-term extrapolation. In addition to the fundamental lack of determinism in quantum physics — the nondeterministic link between the state and the phenomena — there is then a classical-type vagueness of the probabilistic predictions, rather similar to the situation of classical deterministic chaos.

Chapter 2

Kinematics: How Quantum Systems are Described

2.1 Stern–Gerlach experiment

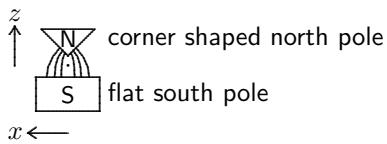
Now, turning to the systematic development of basic concepts and, at the same time, of essential pieces of the mathematical formalism, let us consider the historical *Stern**–*Gerlach*† *experiment* of 1922. In a schematic description of this experiment,



(2.1.1)

silver atoms emerge from an oven (on the right), pass through collimating apertures, thereby forming a well-defined beam of atoms, and then pass through an inhomogeneous magnetic field, eventually reaching a screen where they are collected. The inhomogeneous field is stronger at the top ($z > 0$ side) than the bottom ($z < 0$ side), which is perhaps best seen in a frontal view:

atom beam mainly
at the center;
we are looking at
the approaching
atoms



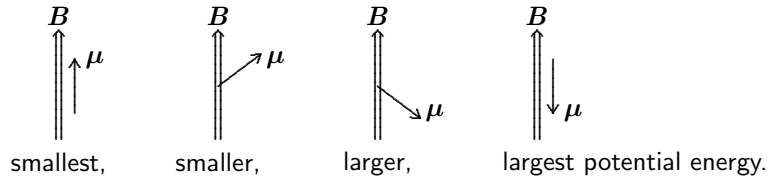
(2.1.2)

*Otto Stern (1888–1969) †Walther Gerlach (1889–1979)

Silver atoms are endowed with a permanent magnetic dipole moment so that there is a potential energy

$$E_{\text{magn}} = -\boldsymbol{\mu} \cdot \boldsymbol{B} \tag{2.1.3}$$

associated with dipole $\boldsymbol{\mu}$ in the magnetic field $\boldsymbol{B}$. It is the smallest if $\boldsymbol{\mu}$ and $\boldsymbol{B}$ are parallel and the largest when they are antiparallel:



(2.1.4)

As a consequence, there is a torque $\boldsymbol{\tau}$ on the dipole exerted by the magnetic field,

$$\boldsymbol{\tau} = \boldsymbol{\mu} \times \boldsymbol{B}\,, \tag{2.1.5}$$

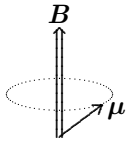
that tends to turn the dipole parallel to $\boldsymbol{B}$. But since the intrinsic angular momentum, the *spin*, of the atom is proportional to $\boldsymbol{\mu}$, and the rate of change of the angular momentum is just the torque, we have

$$\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{\mu} \propto \boldsymbol{\mu} \times \boldsymbol{B}\,, \tag{2.1.6}$$

which is to say that the dipole moment $\boldsymbol{\mu}$ precesses around the direction of $\boldsymbol{B}$, whereby the value of E_{magn} remains unchanged,

$$\frac{\mathrm{d}}{\mathrm{d}t}E_{\text{magn}} = -\frac{\mathrm{d}\boldsymbol{\mu}}{\mathrm{d}t} \cdot \boldsymbol{B} - \boldsymbol{\mu} \cdot \underbrace{\frac{\mathrm{d}\boldsymbol{B}}{\mathrm{d}t}}_{=0} \propto (\boldsymbol{\mu} \times \boldsymbol{B}) \cdot \boldsymbol{B} = 0\,, \tag{2.1.7}$$

pictorially:



(2.1.8)

In a homogeneous magnetic field, this is all that would be happening, but the field of the Stern–Gerlach magnet is inhomogeneous: its strength de-

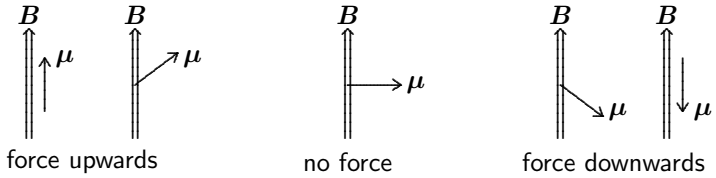
pends on position, mainly growing with increasing z. Therefore, the magnetic energy is position dependent,

$$E_{\text{magn}}(\boldsymbol{r}) = -\boldsymbol{\mu} \cdot \boldsymbol{B}(\boldsymbol{r})\,, \tag{2.1.9}$$

and that gives rise to a force $\boldsymbol{F}$ on the atom, equal to the negative gradient of this magnetic energy,

$$\boldsymbol{F} = -\boldsymbol{\nabla} E_{\text{magn}} = \boldsymbol{\nabla}\big(\boldsymbol{\mu} \cdot \boldsymbol{B}(\boldsymbol{r})\big)\,. \tag{2.1.10}$$

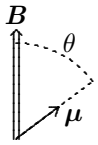
For $\boldsymbol{\mu} \cdot \boldsymbol{B} > 0$, the force is toward the region of stronger $\boldsymbol{B}$ field and for $\boldsymbol{\mu} \cdot \boldsymbol{B} < 0$, it is toward the region of weaker field. Thus,



(2.1.11)

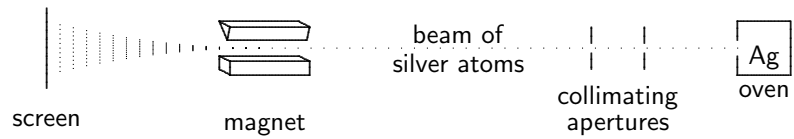
with the strength of the force proportional to the cosine of the angle between $\boldsymbol{\mu}$ and $\boldsymbol{B}$ because

$$\boldsymbol{\mu} \cdot \boldsymbol{B} = |\boldsymbol{\mu}|\,|\boldsymbol{B}|\cos(\theta) \tag{2.1.12}$$



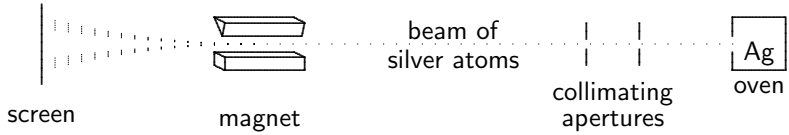
of course.

Now, the atoms emerging from the oven are completely unbiased in their magnetic properties; there is nothing that would prefer one orientation of $\boldsymbol{\mu}$ and discriminate against others. All orientations are equally likely which is to say that they will occur with equal frequency. Thus, some atoms will experience a strong force upwards and others a weaker one, yet others experience forces pulling downwards with a variety of strengths. Clearly, then, we expect the beam of atoms to be spread out over the screen:



(2.1.13)

But this is *not* what is observed in such an experiment and was observed by Stern and Gerlach in 1922. Rather, one finds just two spots:



(2.1.14)

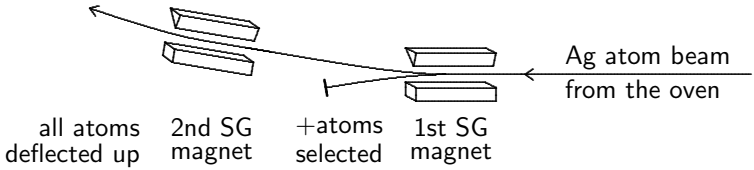
It is *as if* the atoms were prealigned with the magnetic field, some having the dipole moment parallel and some having it antiparallel to $\boldsymbol{B}$. But, of course, there is no such prealignment, nothing prepares the atoms for this particular geometry of the magnetic field. For, just as well, we could have chosen the dominant field component in the x direction. Then, the splitting into two would be along x, perpendicular to z, and the situation would be *as if* the atoms were prealigned in the x direction. Clearly, the assumption of prophetic prealignment is ludicrous.

Rather, we have to accept that the classical prediction of a spread-out beam is inconsistent with the experimental observation. Classical physics fails here, it cannot account for the outcome of the Stern–Gerlach experiment.

Since there is nothing in the preparation of the beam that could possibly bias the atoms toward a particular direction, we expect correctly that half of the atoms are deflected up and half are deflected down. But an individual atom is not split into two, an individual atom is either deflected up or deflected down. And this happens in the perfectly probabilistic manner discussed above. There is no possibility of predicting the fate of the next atom if we think of performing the experiment in a one-atom-at-a-time fashion.

2.2 Successive Stern–Gerlach measurements

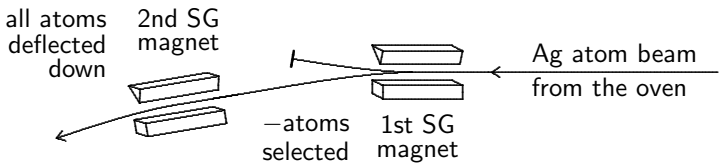
Let us put a label on the atoms: We speak of a +atom when it is deflected up (region $z > 0$) and of a −atom when it is deflected down (region $z < 0$). In view of what we just noted, namely that we cannot possibly predict if an atom will be deflected up or down, it may not be possible to attach such labels. But, in fact, there is a well-defined sense to it. Consider atoms that have been deflected up by the first Stern–Gerlach magnet and are then passed through the second one:



(2.2.1)

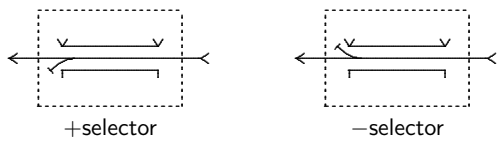
One observes (and perhaps anticipates) that all such atoms are deflected up. That is: +atoms have the objective property that they are predictably deflected up and never down.

Likewise, all −atoms are predictably deflected down:



(2.2.2)

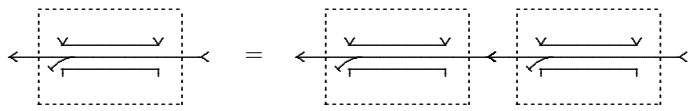
Let us simplify the drawings by a reasonable amount of abstraction:



(2.2.3)

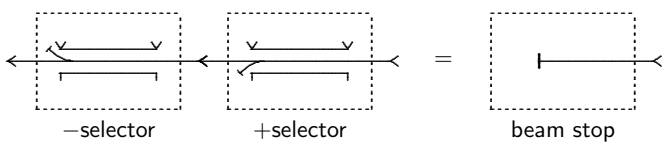
We are going to find mathematical symbols for these operations, but to do this consistently, we must first establish some basic properties.

First, note that a repeated selection of the same kind is equivalent to a single selection, for example,



(2.2.4)

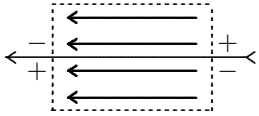
whereas a +selection followed by a −selection blocks the beam completely:



(2.2.5)

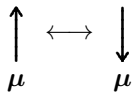
and the same is true for a −selection followed by a +selection.

We shall also need the effect of a homogeneous magnetic field that the atom may pass through:



(2.2.6)

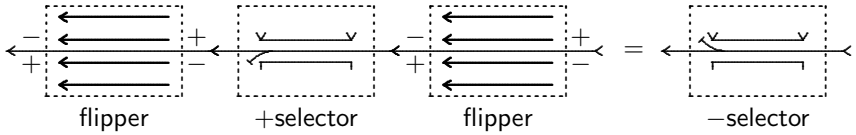
Let us take this field in the y direction and choose its strength such that the precession of the atomic magnetic moment around the y axis amounts to a rotation by 180°:



(2.2.7)

Thus, a flipper turns +atoms into −atoms and −atoms into +atoms.

The sequence flipper, +selector, flipper is equivalent to a −selector:



(2.2.8)

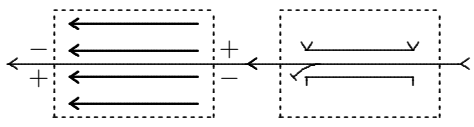
+atoms that enter the first flipper leave it as −atoms and are then rejected by the +selector, whereas −atoms that enter the first flipper leave it as +atoms and the +selector lets them pass through to the second flipper that then turns them back into −atoms. In summary, +atoms are rejected and −atoms are let through. Likewise, we have

$$(\text{flipper})\,(-\text{selector})\,(\text{flipper}) = (+\text{selector})\,, \tag{2.2.9}$$

when the two flippers sandwich a −selector.

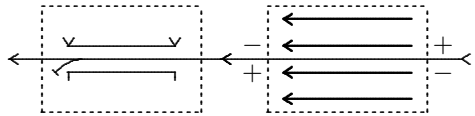
2.3 Order matters

We learn something important by comparing the two orders in which we can have a single flipper and a +selector. A +selector followed by a flipper:



(2.3.1)

is an apparatus that accepts +atoms and turns them into −atoms; the apparatus of the reverse order, a flipper followed by a +selector:



(2.3.2)

accepts −atoms and turns them into +atoms. Clearly, the order matters: the two composed apparatus are very different in their net effect on the atoms.

2.4 Mathematization

This observation tells us that we need a mathematical formalism in which the apparatus (+selector, −selector, flipper, ...) are represented by symbols that can be composed (to make new apparatus) but with a composition law that is not commutative. So, clearly representing apparatus by ordinary numbers with addition or multiplication as the composition law cannot be adequate because the addition and multiplication of numbers *are* commutative — the order does *not* matter.

The simplest mathematical objects with a noncommutative composition law are square matrices and their multiplication. Presently, all we need are 2×2 matrices, for which

$$BA = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = \begin{pmatrix} b_{11}a_{11} + b_{12}a_{21} & b_{11}a_{12} + b_{12}a_{22} \\ b_{21}a_{11} + b_{22}a_{21} & b_{21}a_{12} + b_{22}a_{22} \end{pmatrix} \tag{2.4.1}$$

and

$$AB = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} = \begin{pmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} \\ a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} \end{pmatrix} \tag{2.4.2}$$

are different as a rule: $BA \neq AB$ (important exceptions aside).

And for the description of the state of the atom, fitting to the 2×2 matrices for the apparatus, we shall use two-component columns. The basic ones are

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix} \text{ for +atoms and } \begin{pmatrix} 0 \\ 1 \end{pmatrix} \text{ for −atoms.} \tag{2.4.3}$$

We shall immediately address an important issue: What meaning shall we give to columns such as

$$\begin{pmatrix} -1 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} \mathrm{i} \\ 0 \end{pmatrix}, \quad \text{or more generally } \begin{pmatrix} \alpha \\ 0 \end{pmatrix}? \tag{2.4.4}$$

Except for the ± property — we recall that it is the objective property of being predictably deflected up or down — there is no other property of the atom to be specified. Therefore, the only plausible option is to make no difference between the columns. *All* columns of the form $\begin{pmatrix} \alpha \\ 0 \end{pmatrix}$ are equivalent. And yet we do impose one restriction: The complex number α is normalized to unit modulus, $|\alpha|^2 = 1$. Columns such as $\begin{pmatrix} 2 \\ 0 \end{pmatrix}$ would need to be normalized, but that is a simple matter.

Likewise, all columns of the form $\begin{pmatrix} 0 \\ \beta \end{pmatrix}$ with $|\beta|^2 = 1$ are equivalent; we shall make no effort in distinguishing them from each other. Every column of this kind symbolizes a −atom.

The 2×2 matrix for the +selector is then

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \equiv S_+ \tag{2.4.5}$$

and we verify that it does what it should: it lets +atoms pass unaffected, but rejects −atoms,

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} = 0\,. \tag{2.4.6}$$

Likewise,

$$\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \equiv S_- \tag{2.4.7}$$

will symbolize the −selector, and

$$\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = 0\,, \qquad \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{2.4.8}$$

show that the assignment is consistent.

The flipper interchanges $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ so that $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ is an obvious choice. However, to be later in agreement with standard conventions, we

will use

$$\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \equiv F \tag{2.4.9}$$

instead so that

$$F\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ -1 \end{pmatrix}, \qquad F\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}. \tag{2.4.10}$$

We have already noted that there is no significant difference between $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ and $\begin{pmatrix} 0 \\ -1 \end{pmatrix}$, both stand for "−atom" on equal footing, and so this conventional minus sign in F is of no deeper physical significance.

We put the symbols to a couple of tests, checking if the assignments agree with what we have found earlier. First, two +selections in succession must be like just one,

$$S_+S_+ = S_+ : \quad \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \tag{2.4.11}$$

likewise, for successive −selections,

$$S_-S_- = S_- : \quad \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \tag{2.4.12}$$

and a +selection followed by a −selection must leave nothing,

$$S_-S_+ = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} = 0\,, \tag{2.4.13}$$

indeed. The same is true for the reversed order,

$$S_+S_- = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} = 0\,. \tag{2.4.14}$$

For notational simplicity, we are using the same symbol for the number 0, for the null column in (2.4.6) and (2.4.8), and for the null matrix here.

Note that, keeping in mind that eventually a column stands on the right, these products are such that the earlier operation stands to the right of the later one, as in

$$\underset{\substack{\uparrow \\ \text{+atom entering}}}{S_-}\underset{\substack{\uparrow \\ \text{nothing goes through}}}{S_+}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = S_-\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} = 0\,. \tag{2.4.15}$$

Slightly more involved is the sequence

$$(\text{flipper})(-\text{selector})(\text{flipper}) = (+\text{selector})\,, \tag{2.4.16}$$

where we should verify that

$$FS_+F \quad \text{has the same effect as} \quad S_-\,. \tag{2.4.17}$$

Let us see:

$$\begin{aligned} FS_+F &= \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \\ &= \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & -1 \end{pmatrix} = -S_-\,. \end{aligned} \tag{2.4.18}$$

Indeed, it works, once more with an overall minus sign that must not bother us. Similarly, we should have

$$FS_-F = S_+ \tag{2.4.19}$$

except perhaps for another minus sign of this sort. See

$$\begin{aligned} FS_-F &= \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \\ &= \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\begin{pmatrix} 0 & 0 \\ -1 & 0 \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & 0 \end{pmatrix} = -S_+ \end{aligned} \tag{2.4.20}$$

indeed, and yes there is another minus sign appearing.

Finally, let us see how the two situations of Section 2.3 differ. In (2.3.1) and (2.3.2), we have

$$\begin{aligned} FS_+ &= \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ -1 & 0 \end{pmatrix} \\ \text{and} \quad S_+F &= \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \end{aligned} \tag{2.4.21}$$

respectively. They are very different as they should be, and these situations really represent the physics correctly. We verify that FS_+ accepts +atoms and turns them into −atoms while rejecting −atoms,

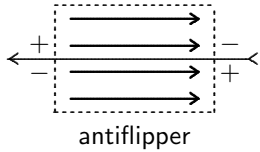
$$\begin{pmatrix} 0 & 0 \\ -1 & 0 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ -1 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 \\ -1 & 0 \end{pmatrix}\begin{pmatrix} 0 \\ 1 \end{pmatrix} = 0\,, \tag{2.4.22}$$

and that S_+F rejects +atoms and turns −atoms into + atoms,

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = 0\,, \quad \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \tag{2.4.23}$$

indeed.

When we first introduced the flipper in (2.2.6), we took the magnetic field pointing in the $+y$ direction. Just as well, we could have it pointing in the $-y$ direction, thus getting the "antiflipper,"



$$F^{-1} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}. \tag{2.4.24}$$

Since it undoes the effect of the flipper — one turning magnetic moments clockwise and the other counter-clockwise — we use the inverse F^{-1} of the flipper matrix F to symbolize the antiflipper. One verifies easily that

$$\begin{aligned} F^{-1}F &= \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = 1\,, \\ FF^{-1} &= \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = 1\,, \end{aligned} \tag{2.4.25}$$

which translate into the physical statement that flipper and antiflipper compensate for each other, irrespective of the order in which they act. Similar to the multiple meanings of the symbol 0 that we pointed out after (2.4.14), here we note that we use the same symbol for the number 1 and for the unit matrix.

The commutativity observed in (2.4.25) is generally true; see Exercise 3. If X, Y, and Z are $n \times n$ square matrices such that $XY = 1$ and $YZ = 1$, then $X = Z$, and we write $X = Z = Y^{-1}$. In other words, there is a *unique inverse*, not one inverse for multiplication on the left and another for multiplication on the right; see Exercise 3.

The matrices F, F^{-1} that we found for the flipper and antiflipper should be particular cases of the matrices that stand for stretches of magnetic field, pointing in the y direction but with varying strength. A rather weak field will not turn +atoms into −atoms but will have some small effect on them, rotating the magnetic moments by some small angle. More generally, we characterize any such field by the angle ϕ by which it rotates the magnetic

moments. Then,

$$R_y(\phi) = \begin{pmatrix} \cos\left(\frac{1}{2}\phi\right) & \sin\left(\frac{1}{2}\phi\right) \\ -\sin\left(\frac{1}{2}\phi\right) & \cos\left(\frac{1}{2}\phi\right) \end{pmatrix} \tag{2.4.26}$$

is appropriate. We check the consistency of

$$\begin{aligned} R_y(\phi = 0) &= 1\,, \\ R_y(\phi = \pi) &= F\,, \\ R_y(\phi = -\pi) &= F^{-1} \end{aligned} \tag{2.4.27}$$

by inspection and verify that two successive rotations of this sort are just one rotation by the net angle,

$$\underbrace{R_y(\phi_2)}_{\substack{\text{then rotate} \\ \text{by angle } \phi_2}} \underbrace{R_y(\phi_1)}_{\substack{\text{first rotate} \\ \text{by angle } \phi_1}} = \underbrace{R_y(\phi_1 + \phi_2)}_{\substack{\text{which is as much as} \\ \text{rotating by angle } \phi_1 + \phi_2}}. \tag{2.4.28}$$

Written out, this is

$$\begin{aligned} &\begin{pmatrix} \cos\left(\frac{1}{2}\phi_2\right) & \sin\left(\frac{1}{2}\phi_2\right) \\ -\sin\left(\frac{1}{2}\phi_2\right) & \cos\left(\frac{1}{2}\phi_2\right) \end{pmatrix} \begin{pmatrix} \cos\left(\frac{1}{2}\phi_1\right) & \sin\left(\frac{1}{2}\phi_1\right) \\ -\sin\left(\frac{1}{2}\phi_1\right) & \cos\left(\frac{1}{2}\phi_1\right) \end{pmatrix} \\ &= \begin{pmatrix} \cos\left(\frac{1}{2}(\phi_1 + \phi_2)\right) & \sin\left(\frac{1}{2}(\phi_1 + \phi_2)\right) \\ -\sin\left(\frac{1}{2}(\phi_1 + \phi_2)\right) & \cos\left(\frac{1}{2}(\phi_1 + \phi_2)\right) \end{pmatrix}, \end{aligned} \tag{2.4.29}$$

where the familiar trigonometric addition theorems

$$\begin{aligned} \cos\left(\tfrac{1}{2}\phi_2\right)\cos\left(\tfrac{1}{2}\phi_1\right) - \sin\left(\tfrac{1}{2}\phi_2\right)\sin\left(\tfrac{1}{2}\phi_1\right) &= \cos\left(\tfrac{1}{2}(\phi_2 + \phi_1)\right), \\ \sin\left(\tfrac{1}{2}\phi_2\right)\cos\left(\tfrac{1}{2}\phi_1\right) + \cos\left(\tfrac{1}{2}\phi_2\right)\sin\left(\tfrac{1}{2}\phi_1\right) &= \sin\left(\tfrac{1}{2}(\phi_2 + \phi_1)\right) \end{aligned} \tag{2.4.30}$$

are applied.

2.5 Probabilities and probability amplitudes

By looking at the effect of a rotation by $\frac{\pi}{2} = 90°$, we can now identify the columns that stand for "+atoms in the x direction" and "−atom in the x direction,"

$$\underbrace{R_y\Big(\tfrac{1}{2}\pi\Big)}_{\substack{\text{rotate}\\ \text{by } 90^\circ}}\underbrace{\begin{pmatrix}1\\0\end{pmatrix}}_{\substack{+\text{ in the } z\\ \text{direction}}} = \frac{1}{\sqrt{2}}\begin{pmatrix}1&1\\-1&1\end{pmatrix}\begin{pmatrix}1\\0\end{pmatrix} = \underbrace{\frac{1}{\sqrt{2}}\begin{pmatrix}1\\-1\end{pmatrix}}_{\substack{-\text{ in the } x\\ \text{direction}}},$$

$$\underbrace{R_y\Big(\tfrac{1}{2}\pi\Big)}_{\substack{\text{rotate}\\ \text{by } 90^\circ}}\underbrace{\begin{pmatrix}0\\1\end{pmatrix}}_{\substack{-\text{ in the } z\\ \text{direction}}} = \frac{1}{\sqrt{2}}\begin{pmatrix}1&1\\-1&1\end{pmatrix}\begin{pmatrix}0\\1\end{pmatrix} = \underbrace{\frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\end{pmatrix}}_{\substack{+\text{ in the } x\\ \text{direction}}}. \tag{2.5.1}$$

If a Stern–Gerlach measurement is performed with the dominant component of the magnetic field in the z direction, atoms of the "$\pm$ in x" kind will be deflected up or down with 50% probability each because we get the "$\pm$ in x" kind halfway between "$+$ in z" and "$-$ in z,"

$$\underbrace{-\begin{pmatrix}0\\1\end{pmatrix}}_{-\text{ in } z} = R_y(\pi)\underbrace{\begin{pmatrix}1\\0\end{pmatrix}}_{+\text{ in } z} = R_y\Big(\tfrac{1}{2}\pi\Big)^2\begin{pmatrix}1\\0\end{pmatrix}$$

$$= R_y\Big(\tfrac{1}{2}\pi\Big)\underbrace{R_y\Big(\tfrac{1}{2}\pi\Big)\begin{pmatrix}1\\0\end{pmatrix}}_{\text{"halfway"}}. \tag{2.5.2}$$

Asking for the probabilities of $+$deflection and $-$deflection, we begin with the ratio 100%:0% for $\begin{pmatrix}1\\0\end{pmatrix}$ and have the ratio 50%:50% halfway in between for $R_y\Big(\frac{1}{2}\pi\Big)\begin{pmatrix}1\\0\end{pmatrix} \mathrel{\widehat{=}}$ "$-$ in x."

Likewise,

$$\begin{pmatrix}0\\1\end{pmatrix} = R_y(-\pi)\begin{pmatrix}1\\0\end{pmatrix} = R_y\Big(-\tfrac{1}{2}\pi\Big)\underbrace{R_y\Big(-\tfrac{1}{2}\pi\Big)\begin{pmatrix}1\\0\end{pmatrix}}_{\text{"halfway"}}$$

$$= R_y\Big(-\tfrac{1}{2}\pi\Big)\frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\end{pmatrix} \tag{2.5.3}$$

says the same thing about $R_y\Big(-\frac{1}{2}\pi\Big)\begin{pmatrix}1\\0\end{pmatrix} \mathrel{\widehat{=}}$ "$+$ in x." Note that

$$\text{"}\pm\text{ in } x\text{"} \mathrel{\widehat{=}} \frac{1}{\sqrt{2}}\begin{pmatrix}1\\\pm1\end{pmatrix} = \frac{1}{\sqrt{2}}\begin{pmatrix}0\\1\end{pmatrix} + \frac{1}{\sqrt{2}}\begin{pmatrix}0\\\pm1\end{pmatrix} \tag{2.5.4}$$

so that the "± in x" states are weighted sums of the "± in z" states, with coefficients whose *squares* are those 50% probabilities,

$$\left|\frac{1}{\sqrt{2}}\right|^2 = \frac{1}{2}\,; \tag{2.5.5}$$

see also Exercise 4.

More generally, we have as the result of a rotation by ϕ

$$R_y(\phi)\begin{pmatrix}1\\0\end{pmatrix} = \begin{pmatrix}\cos\left(\frac{1}{2}\phi\right)\\-\sin\left(\frac{1}{2}\phi\right)\end{pmatrix} = \cos\left(\frac{1}{2}\phi\right)\begin{pmatrix}1\\0\end{pmatrix} + \sin\left(\frac{1}{2}\phi\right)\begin{pmatrix}0\\-1\end{pmatrix} \tag{2.5.6}$$

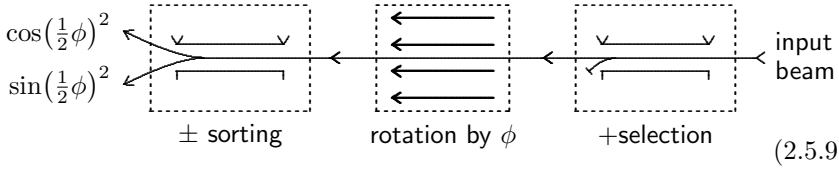
so that $\begin{pmatrix}1\\0\end{pmatrix}$ and $\begin{pmatrix}0\\-1\end{pmatrix}$ are summed with weights $\cos\left(\frac{1}{2}\phi\right)$ and $\sin\left(\frac{1}{2}\phi\right)$, whose *squares* add up to unity,

$$\cos\left(\frac{1}{2}\phi\right)^2 + \sin\left(\frac{1}{2}\phi\right)^2 = 1\,. \tag{2.5.7}$$

This invites the following interpretation:

> The fraction $\cos\left(\frac{1}{2}\phi\right)^2 = \frac{1}{2}\big(1+\cos(\phi)\big)$ of these atoms will be deflected in the $+z$ direction, and the fraction $\sin\left(\frac{1}{2}\phi\right)^2 = \frac{1}{2}\big(1-\cos(\phi)\big)$ will be deflected downwards. (2.5.8)

We can regard the statements as *probabilistic predictions* of our still underdeveloped formalism. We should set out and verify them by the appropriate experiment, namely



(2.5.9)

and such an experiment would confirm the predictions.

Having thus identified columns for "± in z" and "± in x," how about "± in y"? Atoms that have their magnetic moments aligned with the y axis, be it parallel or antiparallel, are not rotated by the flipper at all so that

$$F\begin{pmatrix}\alpha\\\beta\end{pmatrix} = \lambda\begin{pmatrix}\alpha\\\beta\end{pmatrix} \quad \text{if } \begin{pmatrix}\alpha\\\beta\end{pmatrix} \text{ is "+ in } y\text{" or "− in } y\text{"} \tag{2.5.10}$$

where λ is any complex number. With $F = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$, this says

$$\left.\begin{aligned} \beta &= \lambda\alpha \\ -\alpha &= \lambda\beta \end{aligned}\right\} \quad \text{or} \quad -\alpha\beta = \lambda^2\alpha\beta\,. \tag{2.5.11}$$

The solution $\alpha = \beta = 0$ is not permitted because is would give us the null column $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$, so we must have

$$\lambda = +\mathrm{i} \quad \text{or} \quad \lambda = -\mathrm{i}\,. \tag{2.5.12}$$

In the first case, we get

$$\frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ \mathrm{i} \end{pmatrix} \mathrel{\widehat{=}} \text{“}+ \text{ in } y\text{”} \tag{2.5.13}$$

and

$$\frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -\mathrm{i} \end{pmatrix} \mathrel{\widehat{=}} \text{“}- \text{ in } y\text{”} \tag{2.5.14}$$

results in the second case. The factors of $\frac{1}{\sqrt{2}}$ are supplied in analogy with the x states so that *squaring* the prefactors in

$$\frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ \pm\mathrm{i} \end{pmatrix} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 0 \end{pmatrix} \pm \frac{\mathrm{i}}{\sqrt{2}}\begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{2.5.15}$$

again gives the probabilities of

$$\left|\frac{1}{\sqrt{2}}\right|^2 = \frac{1}{2} \quad \text{and} \quad \left|\pm\frac{\mathrm{i}}{\sqrt{2}}\right|^2 = \frac{1}{2} \tag{2.5.16}$$

for deflection up or down in a z-deflection Stern–Gerlach measurement.

Consistency requires that all spatial directions are on equal footing. So, atoms prepared (= preselected) as “$-$ in y” or “$+$ in y” should also be deflected 50%:50% in an x deflection measurement. According to the rules just established, we verify this by writing the “$\pm$ in y” states as weighted sums of the “$\pm$ in x” states and then *square* the coefficients. Let us see,

$$\underbrace{\frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ \pm\mathrm{i} \end{pmatrix}}_{\pm \text{ in } y} = \frac{1 \pm \mathrm{i}}{2}\,\underbrace{\frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix}}_{+ \text{ in } x} + \frac{1 \mp \mathrm{i}}{2}\,\underbrace{\frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix}}_{- \text{ in } x}, \tag{2.5.17}$$

so the probabilities in question are

$$\left|\frac{1 \pm \mathrm{i}}{2}\right|^2 = \frac{(1 \mp \mathrm{i})(1 \pm \mathrm{i})}{4} = \frac{1}{2}\,, \tag{2.5.18}$$

indeed.

Consistent with (2.5.4) and (2.5.5) as well as (2.5.15) and (2.5.16), and more general than (2.5.6) and (2.5.7), we normalize an arbitrary column $\begin{pmatrix} \alpha \\ \beta \end{pmatrix}$ by

$$|\alpha|^2 + |\beta|^2 = 1\,, \tag{2.5.19}$$

with the ultimate justification given shortly when we arrive at (2.5.40). How do we accomplish the presentation of such a column in terms of the "$\pm$ in x," or "$\pm$ in y," or "$\pm$ in z" columns quite generally? It is easy for z,

$$\begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \alpha \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \beta \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{2.5.20}$$

with coefficients

$$\alpha = \begin{pmatrix} 1 & 0 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \tag{2.5.21}$$

and

$$\beta = \begin{pmatrix} 0 & 1 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix}. \tag{2.5.22}$$

For the x columns, we have

$$\begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \frac{\alpha + \beta}{\sqrt{2}} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \frac{\alpha - \beta}{\sqrt{2}} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix} \tag{2.5.23}$$

with coefficients

$$\frac{\alpha + \beta}{\sqrt{2}} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \tag{2.5.24}$$

and

$$\frac{\alpha - \beta}{\sqrt{2}} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix}. \tag{2.5.25}$$

Finally, for the "$\pm$ in y" columns, we have

$$\begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \frac{\alpha - \mathrm{i}\beta}{\sqrt{2}} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ \mathrm{i} \end{pmatrix} + \frac{\alpha + \mathrm{i}\beta}{\sqrt{2}} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -\mathrm{i} \end{pmatrix} \tag{2.5.26}$$

with coefficients

$$\frac{\alpha - \mathrm{i}\beta}{\sqrt{2}} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & -\mathrm{i} \end{pmatrix}\begin{pmatrix} \alpha \\ \beta \end{pmatrix} \tag{2.5.27}$$

and

$$\frac{\alpha + \mathrm{i}\beta}{\sqrt{2}} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & \mathrm{i} \end{pmatrix}\begin{pmatrix} \alpha \\ \beta \end{pmatrix}. \tag{2.5.28}$$

Clearly, there is a general pattern here, namely the coefficients of one particular decomposition are obtained by multiplying $\begin{pmatrix} \alpha \\ \beta \end{pmatrix}$ with the corresponding *rows*, obtained by taking the adjoints of the *columns* in questions, that is, the complex conjugate of the transposed columns. It is important here that the columns we use for the weighted sum are orthogonal to each other. In general, then, we have this. There are two columns

$$\begin{pmatrix} a_1 \\ b_1 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} a_2 \\ b_2 \end{pmatrix} \tag{2.5.29}$$

that are normalized

$$|a_1|^2 + |b_1|^2 = 1\,, \qquad |a_2|^2 + |b_2|^2 = 1 \tag{2.5.30}$$

and orthogonal to each other,

$$0 = \begin{pmatrix} a_1 \\ b_1 \end{pmatrix}^{\dagger}\begin{pmatrix} a_2 \\ b_2 \end{pmatrix} = \begin{pmatrix} a_1^* & b_1^* \end{pmatrix}\begin{pmatrix} a_2 \\ b_2 \end{pmatrix} = a_1^* a_2 + b_1^* b_2 \tag{2.5.31}$$

or, equivalently after complex conjugation,

$$0 = \begin{pmatrix} a_2 \\ b_2 \end{pmatrix}^{\dagger}\begin{pmatrix} a_1 \\ b_1 \end{pmatrix} = \begin{pmatrix} a_2^* & b_2^* \end{pmatrix}\begin{pmatrix} a_1 \\ b_1 \end{pmatrix} = a_2^* a_1 + b_2^* b_1\,, \tag{2.5.32}$$

where the dagger † indicates the adjoint. Then,

$$\begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} a_1 \\ b_1 \end{pmatrix}\begin{pmatrix} a_1^* & b_1^* \end{pmatrix}\begin{pmatrix} \alpha \\ \beta \end{pmatrix} + \begin{pmatrix} a_2 \\ b_2 \end{pmatrix}\begin{pmatrix} a_2^* & b_2^* \end{pmatrix}\begin{pmatrix} \alpha \\ \beta \end{pmatrix} \tag{2.5.33}$$

as we can verify by multiplying from the left by $\begin{pmatrix} a_1^* & b_1^* \end{pmatrix} = \begin{pmatrix} a_1 \\ b_1 \end{pmatrix}^{\dagger}$ and $\begin{pmatrix} a_2^* & b_2^* \end{pmatrix} = \begin{pmatrix} a_2 \\ b_2 \end{pmatrix}^{\dagger}$, respectively. All particular decompositions in (2.5.20)–(2.5.28) are examples for this general rule, for the pairs of columns that stand for "± in z," "± in x," and "± in y."

But there is yet another way of looking at this. So far we have read

$$\underbrace{\begin{pmatrix} a_1 \\ b_1 \end{pmatrix}}_{\substack{\text{column in the} \\ \text{decomposition}}} \underbrace{\begin{pmatrix} a_1^* & b_1^* \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix}}_{\substack{\text{coefficient that goes} \\ \text{with this column}}} . \tag{2.5.34}$$

Now, let us choose another association, which is allowed because matrix multiplication is associative,

$$\underbrace{\begin{pmatrix} a_1 \\ b_1 \end{pmatrix} \begin{pmatrix} a_1^* & b_1^* \end{pmatrix}}_{\substack{\text{matrix that projects} \\ \text{on column } \begin{pmatrix} a_1 \\ b_1 \end{pmatrix}}} \underbrace{\begin{pmatrix} \alpha \\ \beta \end{pmatrix}}_{\text{input}} . \tag{2.5.35}$$

The *projection* property is verified by

$$\left(\begin{pmatrix} a_1 \\ b_1 \end{pmatrix} \begin{pmatrix} a_1^* & b_1^* \end{pmatrix} \right) \begin{pmatrix} a_1 \\ b_1 \end{pmatrix} = \begin{pmatrix} a_1 \\ b_1 \end{pmatrix} \underbrace{\left(\begin{pmatrix} a_1^* & b_1^* \end{pmatrix} \begin{pmatrix} a_1 \\ b_1 \end{pmatrix} \right)}_{=1 \text{ by normalization}} = \begin{pmatrix} a_1 \\ b_1 \end{pmatrix} \tag{2.5.36}$$

and

$$\left(\begin{pmatrix} a_1 \\ b_1 \end{pmatrix} \begin{pmatrix} a_1^* & b_1^* \end{pmatrix} \right) \begin{pmatrix} a_2 \\ b_2 \end{pmatrix} = \begin{pmatrix} a_1 \\ b_1 \end{pmatrix} \underbrace{\left(\begin{pmatrix} a_1^* & b_1^* \end{pmatrix} \begin{pmatrix} a_2 \\ b_2 \end{pmatrix} \right)}_{=0 \text{ by orthogonality}} = 0 , \tag{2.5.37}$$

indeed.

And, finally, upon reading the decomposition (2.5.33) as involving the sum of two projectors,

$$\begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \left(\underbrace{\begin{pmatrix} a_1 \\ b_1 \end{pmatrix} \begin{pmatrix} a_1^* & b_1^* \end{pmatrix}}_{\substack{\text{projects on} \\ \text{column } \begin{pmatrix} a_1 \\ b_1 \end{pmatrix}}} + \underbrace{\begin{pmatrix} a_2 \\ b_2 \end{pmatrix} \begin{pmatrix} a_2^* & b_2^* \end{pmatrix}}_{\substack{\text{projects on} \\ \text{column } \begin{pmatrix} a_2 \\ b_2 \end{pmatrix}}} \right) \begin{pmatrix} \alpha \\ \beta \end{pmatrix} , \tag{2.5.38}$$

we note that the sum of the two projection matrices must be the unit matrix,

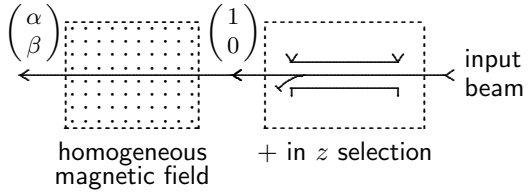
$$\begin{pmatrix} a_1 \\ b_1 \end{pmatrix} \begin{pmatrix} a_1^* & b_1^* \end{pmatrix} + \begin{pmatrix} a_2 \\ b_2 \end{pmatrix} \begin{pmatrix} a_2^* & b_2^* \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = 1 . \tag{2.5.39}$$

Harkening back to (2.5.4)–(2.5.18), we recall that in all particular examples discussed then, *probabilities* for deflection in the respective + and − directions were always correctly given by the absolute squares of the coefficients. This is why the coefficients are termed *probability amplitudes*. They are explicitly stated in (2.5.20)–(2.5.28), and we find the resulting probabilities as

$$\begin{aligned}
&\text{prob}(+ \text{ in } z) = |\alpha|^2 ,\\
&\text{prob}(- \text{ in } z) = |\beta|^2 ,\\
&\text{prob}(\pm \text{ in } x) = \left|\frac{\alpha \pm \beta}{\sqrt{2}}\right|^2 = \frac{1}{2}(|\alpha|^2 + |\beta|^2 \pm \alpha^*\beta \pm \beta^*\alpha)\\
&\qquad\qquad\qquad\qquad = \frac{1}{2} \pm \text{Re}(\alpha^*\beta) ,\\
&\text{prob}(\pm \text{ in } y) = \left|\frac{\alpha \mp \mathrm{i}\beta}{\sqrt{2}}\right|^2 = \frac{1}{2}(|\alpha|^2 + |\beta|^2 \mp \mathrm{i}\alpha^*\beta \pm \mathrm{i}\beta^*\alpha)\\
&\qquad\qquad\qquad\qquad = \frac{1}{2} \pm \text{Im}(\alpha^*\beta) ,
\end{aligned} \tag{2.5.40}$$

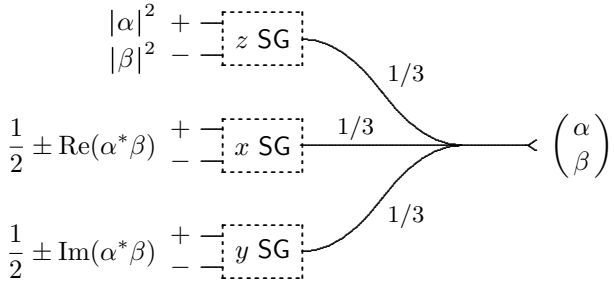
where the normalization $|\alpha|^2 + |\beta|^2 = 1$ of (2.5.19) enters, which we now recognize as the obvious condition that the probabilities for "+ in z" and "− in z" must add up to 100%.

The operational meaning of these statements is as follows. Suppose we have atoms + selected in the z direction and then passed through a homogeneous magnetic field of unknown strength and direction,



(2.5.41)

so that the atoms that emerge are described by some column $\begin{pmatrix}\alpha\\ \beta\end{pmatrix}$, but we do not know the values of α and β. How can we determine them? By performing a z-deflection Stern–Gerlach experiment on a fraction of them (say $\frac{1}{3}$ of all the atoms), an x-Stern–Gerlach measurement on another fraction (another $\frac{1}{3}$, say), and a y-Stern–Gerlach measurement on the rest:



(2.5.42)

Since $|\alpha|^2 + |\beta|^2 = 1$, the z measurement tells us the value of $|\alpha|^2 - |\beta|^2$. It is a number between -1 and $+1$, and we parameterize it conveniently by writing

$$|\alpha|^2 - |\beta|^2 = \cos(2\theta) \quad \text{with} \quad 0 \le \theta \le \frac{\pi}{2}\,. \tag{2.5.43}$$

Then,

$$\begin{aligned} |\alpha|^2 &= \frac{1}{2}\Big(|\alpha|^2 + |\beta|^2\Big) + \frac{1}{2}\Big(|\alpha|^2 - |\beta|^2\Big) \\ &= \frac{1}{2} + \frac{1}{2}\cos(2\theta) = \cos(\theta)^2 \end{aligned} \tag{2.5.44}$$

and

$$\begin{aligned} |\beta|^2 &= \frac{1}{2}\Big(|\alpha|^2 + |\beta|^2\Big)\frac{1}{2}\Big(|\alpha|^2 - |\beta|^2\Big) \\ &= \frac{1}{2} - \frac{1}{2}\cos(2\theta) = \sin(\theta)^2\,. \end{aligned} \tag{2.5.45}$$

Consequently, we have

$$\alpha = \mathrm{e}^{\mathrm{i}\varphi}\cos(\theta)\,, \qquad \beta = \mathrm{e}^{\mathrm{i}\phi}\sin(\theta)\,, \tag{2.5.46}$$

where $\mathrm{e}^{\mathrm{i}\varphi}$, $\mathrm{e}^{\mathrm{i}\phi}$ are as yet undetermined phase factors, complex numbers of unit modulus,

$$\Big|\,\mathrm{e}^{\mathrm{i}\varphi}\Big| = \big|\cos(\varphi) + \mathrm{i}\sin(\varphi)\big| = \sqrt{\cos(\varphi)^2 + \sin(\varphi)^2} = 1\,. \tag{2.5.47}$$

We have here the first of many applications of *Euler's** *identity*

$$\mathrm{e}^{\mathrm{i}\varphi} = \cos(\varphi) + \mathrm{i}\sin(\varphi)\,, \tag{2.5.48}$$

arguably the most famous of his numerous discoveries.

*Leonhard EULER (1707–1783)

Then, taking the values of $\mathrm{Re}(\alpha^*\beta)$ and $\mathrm{Im}(\alpha^*\beta)$ from the x and y measurements, we establish the value of

$$\alpha^*\beta = \mathrm{Re}(\alpha^*\beta) + \mathrm{i}\,\mathrm{Im}(\alpha^*\beta) \tag{2.5.49}$$

and compare with

$$\begin{aligned}\alpha^*\beta &= \big(\mathrm{e}^{\mathrm{i}\varphi}\cos(\theta)\big)^*\big(\mathrm{e}^{\mathrm{i}\phi}\sin(\theta)\big) = \mathrm{e}^{\mathrm{i}(\phi-\varphi)}\cos(\theta)\sin(\theta)\\ &= \frac{1}{2}\,\mathrm{e}^{\mathrm{i}(\phi-\varphi)}\sin(2\theta)\,.\end{aligned} \tag{2.5.50}$$

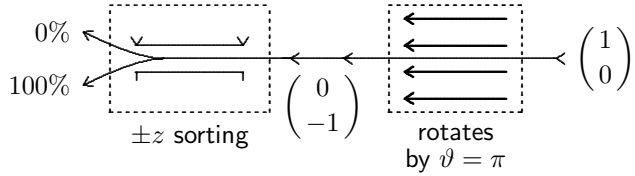
Since we have found $\cos(2\theta)$ and thus $\sin(2\theta) \geq 0$ earlier from the data of the z measurement, we can now determine the value of $\mathrm{e}^{\mathrm{i}(\phi-\varphi)}$ and then know

$$\begin{pmatrix}\alpha\\ \beta\end{pmatrix} = \mathrm{e}^{\mathrm{i}\varphi}\begin{pmatrix}\cos(\theta)\\ \mathrm{e}^{\mathrm{i}(\phi-\varphi)}\sin(\theta)\end{pmatrix} \tag{2.5.51}$$

except for the prefactor of $\mathrm{e}^{\mathrm{i}\varphi}$. But this prefactor is an overall phase factor of no physical significance, so we do not need it. It is consistent with this lack of significance that $\mathrm{e}^{\mathrm{i}\varphi}$ cannot be determined from the experimental data. We can adopt any convenient convention, such as putting $\mathrm{e}^{\mathrm{i}\varphi} = 1$ or $\mathrm{e}^{\mathrm{i}\varphi} = \mathrm{e}^{-\mathrm{i}\frac{1}{2}(\phi-\varphi)}$, for example; all choices are equally good.

2.6 Quantum Zeno effect

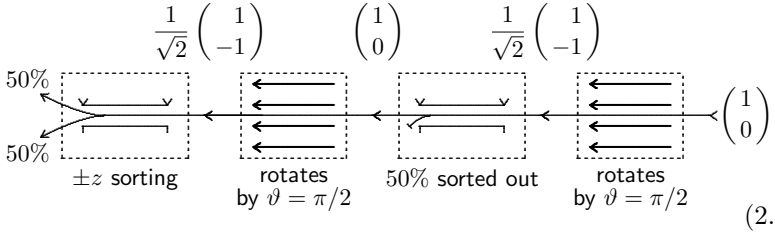
Here is a little application of what we have learned, to a somewhat amusing but also instructive situation. Imagine atoms selected as "+ in z," that is $\begin{pmatrix}1\\0\end{pmatrix}$, and then propagating through a magnetic field that gradually turns them into $\begin{pmatrix}0\\-1\end{pmatrix}$ with $R_y(\vartheta)\begin{pmatrix}1\\0\end{pmatrix} = \begin{pmatrix}\cos\left(\frac{1}{2}\vartheta\right)\\ -\sin\left(\frac{1}{2}\vartheta\right)\end{pmatrix}$ at intermediate stages, eventually followed by a final z measurement:



(2.6.1)

All atoms will be deflected as "− in z" at the end, none as "+ in z." But now let us introduce a control measurement that selects "+ in z" halfway

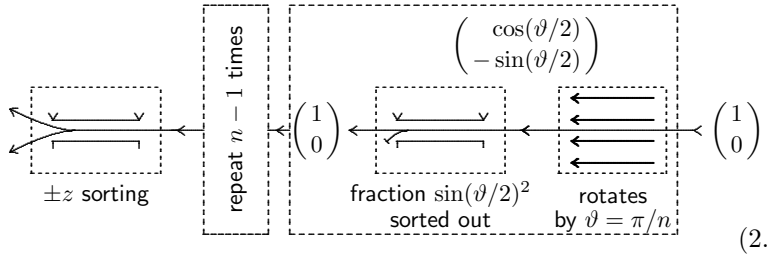
along the way:



(2.6.2)

The probability that an entering atom is eventually deflected as "+ in z" is $50\% \times 50\% = 25\%$ because it has a 50% chance of surviving the midway sorting and another 50% of up-deflection at the end.

As a generalization, let us now break up the magnetic field in n portions, with "+ in z" selectors in between:



(2.6.3)

After passing through a single stretch of the field, we have

$$R_y\Big(\frac{\pi}{n}\Big)\begin{pmatrix}1\\0\end{pmatrix} = \begin{pmatrix}\cos\big(\frac{\pi}{2n}\big)\\-\sin\big(\frac{\pi}{2n}\big)\end{pmatrix} \tag{2.6.4}$$

so that the probability of making it to the next stage is

$$\left|\,(1,0)\begin{pmatrix}\cos\big(\frac{\pi}{2n}\big)\\-\sin\big(\frac{\pi}{2n}\big)\end{pmatrix}\right|^2 = \cos\Big(\frac{\pi}{2n}\Big)^2. \tag{2.6.5}$$

Since there are n stages of this sort all together, we have

$$p_n = \left(\cos\Big(\frac{\pi}{2n}\Big)^2\right)^n = \cos\Big(\frac{\pi}{2n}\Big)^{2n} \tag{2.6.6}$$

for the total probability of getting through. Upon completing Exercise 9, you will see that starting from $p_1 = 0$, $p_2 = \dfrac{1}{4}$, it increases rapidly and approaches unity for $n \to \infty$. Of course, this limit itself is unphysical, but large values of n are not, so let us see what happens for $n \gg 1$, when

$\phi = \dfrac{\pi}{2n}$ is very small and $\cos(\phi) \cong 1 - \dfrac{1}{2}\phi^2$ is a permissible approximation,

$$p_n = \cos\Big(\frac{\pi}{2n}\Big)^{2n} \cong \Big(1 - \frac{1}{2}\Big(\frac{\pi}{2n}\Big)^2\Big)^{2n} \cong \Big(\mathrm{e}^{-\frac{1}{2}\left(\frac{\pi}{2n}\right)^2}\Big)^{2n} = \mathrm{e}^{-\frac{\pi^2}{4n}} \cong 1 - \frac{\pi^2}{4n}\,. \tag{2.6.7}$$

This says that the survival probability is very close to 100% if we only take n sufficiently large; see Exercise 10.

Physically speaking, there is a magnetic field that all by itself would turn "+ in z" $\widehat{=} \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ into "− in z" $\widehat{=} \begin{pmatrix} 0 \\ -1 \end{pmatrix}$, but the repeated checking "is it still $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ or already $\begin{pmatrix} 0 \\ -1 \end{pmatrix}$?" effectively "freezes" the evolution, and for $n \gg 1$, it is as if there were no magnetic field at all. This freezing of the motion is reminiscent of one of the four famous paradoxes invented by Zeno,* and the quantum version just discussed has become known as the *quantum Zeno effect*. It is a real physical phenomenon that is easy to demonstrate if one uses polarized photons rather than magnetic atoms.

But, putting this folklore aside, what we learn from this discussion is that a measurement disturbs the system. The evolution of $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ to $\begin{pmatrix} 0 \\ -1 \end{pmatrix}$ does not proceed as usual if you do not let the atoms alone. Every interference with their evolution disturbs or even disrupts it. You cannot check some properties of an atomic system without introducing a sizeable disturbance. This is quite different from the situation in classical physics where you can measure, say, the temperature of the water in a glass, without altering the water temperature (or any other property) noticeably.

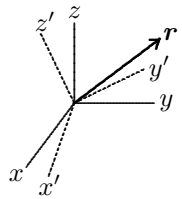
2.7 Kets and bras

Let us return to the statements in (2.5.20), (2.5.23), and (2.5.26),

$$\begin{aligned} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} &= \alpha \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \beta \begin{pmatrix} 0 \\ 1 \end{pmatrix} && (z \text{ columns}) \\ &= \frac{\alpha+\beta}{\sqrt{2}} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \frac{\alpha-\beta}{\sqrt{2}} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix} && (x \text{ columns}) \\ &= \frac{\alpha-\mathrm{i}\beta}{\sqrt{2}} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ \mathrm{i} \end{pmatrix} + \frac{\alpha+\mathrm{i}\beta}{\sqrt{2}} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -\mathrm{i} \end{pmatrix} && (y \text{ columns}), \end{aligned} \tag{2.7.1}$$

*ZENO of Elea (495–430 BC)

which decompose the general column $\begin{pmatrix}\alpha\\ \beta\end{pmatrix}$ in three different ways that refer to "$\pm$ in z," "$\pm$ in x," and "$\pm$ in y," respectively. This has a familiar ring; it reminds us, and quite correctly so, of the representation of one and the same three-dimensional vector by numerical coefficients (sets of three) with respect to a prechosen coordinate system,



$$\begin{aligned}\boldsymbol{r} &= x\,\boldsymbol{e}_x + y\,\boldsymbol{e}_y + z\,\boldsymbol{e}_z\\ &= x'\,\boldsymbol{e}'_x + y'\,\boldsymbol{e}'_y + z'\,\boldsymbol{e}'_z\,.\end{aligned} \tag{2.7.2}$$

While we have one vector $\boldsymbol{r}$ only, its numerical representation,

$$\boldsymbol{r} \mathrel{\widehat{=}} \begin{pmatrix}x\\ y\\ z\end{pmatrix} \quad \text{or} \quad \boldsymbol{r} \mathrel{\widehat{=}} \begin{pmatrix}x'\\ y'\\ z'\end{pmatrix}, \tag{2.7.3}$$

may involve rather different sets of three numbers, depending on which coordinate system, which set of unit vectors they refer to. Thus, it is essential that we do not confuse the vector $\boldsymbol{r}$ with its coordinates. The vector is a geometrical object all by itself, whereas its coordinates also contain reference to the coordinate system, and that we can choose quite arbitrarily, doing whatever is convenient for the calculation.

Matters are quite analogous for the magnetic atoms. The columns $\begin{pmatrix}\alpha\\ \beta\end{pmatrix}$ that we have been using make explicit reference to Stern–Gerlach measurements in what we call the z direction. But that is, of course, just as arbitrary as any such convention. Equally well, we could have singled out the x direction or any other one. Thus, collecting the pairs of probability amplitudes in the three decompositions in (2.7.1), we can say that atoms prepared in a particular manner — as, for example, indicated in (2.5.41) — are equivalently characterized by the columns

$$\begin{pmatrix}\alpha\\ \beta\end{pmatrix}_z, \qquad \frac{1}{\sqrt{2}}\begin{pmatrix}\alpha+\beta\\ \alpha-\beta\end{pmatrix}_x, \qquad \frac{1}{\sqrt{2}}\begin{pmatrix}\alpha-\mathrm{i}\beta\\ \alpha+\mathrm{i}\beta\end{pmatrix}_y, \tag{2.7.4}$$

where the subscript indicates the reference columns that are weighted in (2.7.1).

To free ourselves from the eternal reference to particular basic columns, which are always subject to conventions and never free of arbitrariness,

we take a step that is analogous to going from the numerical component column $\begin{pmatrix} x \\ y \\ z \end{pmatrix}$ to vector $\boldsymbol{r}$. The *state vectors* in the quantum-mechanical formalism are denoted by $|\ \rangle$, with appropriate identifying labels inside, and are called *kets* or ket vectors. As a reminder of the arrows we draw for magnetic moments, we use $\uparrow_z$, $\downarrow_z$ *et cetera* as the identifying labels. Thus,

$$\begin{aligned} &\text{the ket } \big|\uparrow_z\big\rangle \text{ represents an atom that is ``+ in } z\text{''} \\ &\text{and the ket } \big|\downarrow_z\big\rangle \text{ represents an atom that is ``} - \text{ in } z\text{''} \end{aligned} \tag{2.7.5}$$

and likewise, $\big|\uparrow_x\big\rangle$, $\big|\downarrow_x\big\rangle$ for "+ in x," "$-$ in x," as well as $\big|\uparrow_y\big\rangle$, $\big|\downarrow_y\big\rangle$ for atoms that are "$\pm$ in y."

The three decompositions of a general column in (2.7.1) are now presented as

$$\begin{aligned} |\ \rangle = \big|\uparrow_z\big\rangle\alpha + \big|\downarrow_z\big\rangle\beta &= \big|\uparrow_x\big\rangle\frac{\alpha+\beta}{\sqrt{2}} + \big|\downarrow_x\big\rangle\frac{\alpha-\beta}{\sqrt{2}} \\ &= \big|\uparrow_y\big\rangle\frac{\alpha-\mathrm{i}\beta}{\sqrt{2}} + \big|\downarrow_y\big\rangle\frac{\alpha+\mathrm{i}\beta}{\sqrt{2}}\,. \end{aligned} \tag{2.7.6}$$

We are here expressing *one and the same ket* $|\ \rangle$ in terms of different reference kets, fully analogous to $\boldsymbol{r} = \cdots = \cdots$ in (2.7.3).

The two trios of unit vectors in (2.7.2) are two bases in the space of three-dimensional vectors. The analog in (2.7.6) tells us that the three pairs of kets $|\uparrow_z\rangle, |\downarrow_z\rangle$ and $|\uparrow_x\rangle, |\downarrow_x\rangle$ and $|\uparrow_y\rangle, |\downarrow_y\rangle$ are bases for the two-dimensional ket space. The similarity or analogy observed here is mathematical in nature; from the physics point of view, there are substantial differences, namely that the coefficients in (2.7.2) are real and mean distances whereas the coefficients in (2.7.6) are complex probability amplitudes.

Since the abstract kets are numerically represented by columns as we used them so far, they inherit in full their linear vector properties, that is, we can multiply kets with complex numbers to make other kets, and we can form weighted sums of them and get more kets in this manner. The weights that appear in the sums in (2.7.6) continue to have, of course, the physical meaning that we have identified before and noted just now: they are the probability amplitudes that characterize $|\ \rangle$ and tell us, upon squaring, the probabilities of $\uparrow_z$ and $\downarrow_z$, or $\uparrow_x$ and $\downarrow_x$, or $\uparrow_y$ and $\downarrow_y$.

The systematic method for calculating probability amplitudes for given columns $\begin{pmatrix} \alpha \\ \beta \end{pmatrix}$ is, recall the discussion of (2.5.20)–(2.5.33), by multiplying with appropriate rows from the left. These rows are obtained as hermitian

conjugates, or adjoints, of the respective columns. We thus introduce yet another sort of vectors, the adjoints of the kets, denoted by $\langle\ |$ and called *bras*,

$$|\ \rangle^\dagger = \langle\ |\,, \qquad \langle\ |^\dagger = |\ \rangle\,, \tag{2.7.7}$$

and, as the second equation states, the kets are the adjoints of the bras. It follows, perhaps not unexpectedly, that taking the adjoint twice amounts to doing nothing.

Taking the adjoint is essentially a linear operation, but we must remember about the complex conjugation that is involved so that we have

$$\Big(|1\rangle a_1 + |2\rangle a_2 + |3\rangle a_3\Big)^\dagger = a_1^*\langle 1| + a_2^*\langle 2| + a_3^*\langle 3| \tag{2.7.8}$$

for the adjoint of a weighted sum of three kets, for example. More specifically, the adjoint statement to (2.7.6) is

$$\begin{aligned}\langle\ |\ = |\ \rangle^\dagger &= \alpha^*\langle\uparrow_z| + \beta^*\langle\downarrow_z| \\ &= \frac{\alpha^*+\beta^*}{\sqrt{2}}\langle\uparrow_x| + \frac{\alpha^*-\beta^*}{\sqrt{2}}\langle\downarrow_x| \\ &= \frac{\alpha^*+\mathrm{i}\beta^*}{\sqrt{2}}\langle\uparrow_y| + \frac{\alpha^*-\mathrm{i}\beta^*}{\sqrt{2}}\langle\downarrow_y|\,.\end{aligned} \tag{2.7.9}$$

It should be clear that bras and kets are on equal footing. Whatever physical fact can be phrased as a relation among kets can just as well be formulated as the adjoint relation among bras.

By taking the adjoint, we turn the three bases of kets that we recognized in (2.7.6) into the three bases of bras that we see in (2.7.9), namely $\langle\uparrow_z|, \langle\downarrow_z|$ and $\langle\uparrow_x|, \langle\downarrow_x|$ and $\langle\uparrow_y|, \langle\downarrow_y|$. Just like there are many more trios of vectors that can serve as bases for the three-dimensional vectors in (2.7.2), there are many more pairs of kets and their adjoint pairs of bras that are bases for the two-dimensional ket and bra spaces, respectively.

2.8 Brackets, bra-kets, and ket-bras

Looking at, for example, the second line in (2.7.6) and its numerical version in (2.7.1), we note that the probability amplitudes for $\uparrow_x$, $\downarrow_x$ are

$$\frac{\alpha \pm \beta}{\sqrt{2}} = \frac{1}{\sqrt{2}}\begin{pmatrix}1 & \pm 1\end{pmatrix}\begin{pmatrix}\alpha \\ \beta\end{pmatrix}, \tag{2.8.1}$$

which translates into

$$\frac{\alpha+\beta}{\sqrt{2}} = \langle\uparrow_x|\ |\ \rangle = \langle\uparrow_x|\ \rangle\,,$$
$$\frac{\alpha-\beta}{\sqrt{2}} = \langle\downarrow_x|\ |\ \rangle = \langle\downarrow_x|\ \rangle\,, \tag{2.8.2}$$

where we have adopted the usual convention to not write two vertical lines where bra meets ket: $\langle 1|\text{times}|2\rangle = \langle 1|2\rangle$, corresponding to the numerical row-times-column product.

It follows that the normalization of the basis kets to "unit length" is expressed by

$$\langle\uparrow_z|\uparrow_z\rangle = 1\,, \qquad \langle\downarrow_z|\downarrow_z\rangle = 1 \tag{2.8.3}$$

and their orthogonality by

$$\langle\uparrow_z|\downarrow_z\rangle = 0\,, \qquad \langle\downarrow_z|\uparrow_z\rangle = 0\,, \tag{2.8.4}$$

and analogously for the x and y pairs of kets and bras. If you wonder about the origin of the somewhat frivolous terminology, introduced by Dirac,* the observation

$$\underbrace{\langle 1|}_{\text{bra}}\ \underbrace{\times}_{\text{times}}\ \underbrace{|2\rangle}_{\text{ket}} = \underbrace{\langle 1|\ |2\rangle}_{\text{bra-ket}} = \underbrace{\langle 1|2\rangle}_{\text{bracket}} \tag{2.8.5}$$

should tell you.

As indicated by the terms normalization, unit length, and orthogonality, the bracket of a bra and a ket is an inner product of two kets or two bras,

$$\langle 1|2\rangle = \Big(|1\rangle, |2\rangle\Big) \quad \text{or} \quad \langle 1|2\rangle = \Big(\langle 1|, \langle 2|\Big)\,, \tag{2.8.6}$$

where we use the standard (,) notation for an inner product. Indeed, a bracket has all the properties required of an inner product as the corresponding statements about rows and columns confirm,

$$\begin{pmatrix}\alpha_1^* & \beta_1^*\end{pmatrix}\begin{pmatrix}\alpha_2\\ \beta_2\end{pmatrix} = \begin{pmatrix}\alpha_1\\ \beta_1\end{pmatrix}^\dagger\begin{pmatrix}\alpha_2\\ \beta_2\end{pmatrix} = \begin{pmatrix}\alpha_1\\ \beta_1\end{pmatrix}\cdot\begin{pmatrix}\alpha_2\\ \beta_2\end{pmatrix}$$
$$\text{or}\quad \begin{pmatrix}\alpha_1^* & \beta_1^*\end{pmatrix}\begin{pmatrix}\alpha_2\\ \beta_2\end{pmatrix} = \begin{pmatrix}\alpha_1^* & \beta_1^*\end{pmatrix}\begin{pmatrix}\alpha_2^* & \beta_2^*\end{pmatrix}^\dagger = \begin{pmatrix}\alpha_1^* & \beta_1^*\end{pmatrix}\cdot\begin{pmatrix}\alpha_2^* & \beta_2^*\end{pmatrix}, \tag{2.8.7}$$

where we use the dot notation for the inner product as we did in (2.7.2). Now, with this richer vocabulary, we can state that the bases of pairs of kets

*Paul Adrien Maurice Dirac (1902–1984)

or bras that we identified in Section 2.7 are bases composed of orthonormal — that is, normalized to unit length and pairwise orthogonal — kets and bras.

So, we have

$$|\,\rangle = |{\uparrow_z}\rangle\alpha + |{\downarrow_z}\rangle\beta = \Big(\,|{\uparrow_z}\rangle\;|{\downarrow_z}\rangle\,\Big)\begin{pmatrix}\alpha\\ \beta\end{pmatrix}, \tag{2.8.8}$$

where we exhibit the reference kets as a two-component row of kets and the probability amplitudes as a numerical two-component column. Then,

$$\begin{pmatrix}\alpha\\ \beta\end{pmatrix} = \begin{pmatrix}\langle{\uparrow_z}|\,\rangle\\ \langle{\downarrow_z}|\,\rangle\end{pmatrix} = \begin{pmatrix}\langle{\uparrow_z}|\\ \langle{\downarrow_z}|\end{pmatrix}|\,\rangle \tag{2.8.9}$$

expresses that column as the product of a column of bras with the ket in question. Now, putting things together,

$$|\,\rangle = \Big(\,|{\uparrow_z}\rangle\;|{\downarrow_z}\rangle\,\Big)\begin{pmatrix}\langle{\uparrow_z}|\\ \langle{\downarrow_z}|\end{pmatrix}|\,\rangle\,, \tag{2.8.10}$$

we learn that

$$\Big(\,|{\uparrow_z}\rangle\;|{\downarrow_z}\rangle\,\Big)\begin{pmatrix}\langle{\uparrow_z}|\\ \langle{\downarrow_z}|\end{pmatrix} = 1 \tag{2.8.11}$$

because otherwise the statement (2.8.10) cannot be true for all kets $|\,\rangle$. The resulting row times column product,

$$|{\uparrow_z}\rangle\langle{\uparrow_z}| + |{\downarrow_z}\rangle\langle{\downarrow_z}| = 1\,, \tag{2.8.12}$$

contains "ket times bra" structures that are the abstract generalization of "column times row" products that yield matrices. And just like a matrix acting on a column produces a new column, a "ket-bra" applied to a ket gives a new ket,

$$\underbrace{\Big(|1\rangle\langle 2|\Big)}_{\text{ket-bra}}\underbrace{|3\rangle}_{\text{ket}} = \underbrace{|1\rangle}_{\text{ket}}\underbrace{\langle 2|3\rangle}_{\text{bracket}}\,, \tag{2.8.13}$$

and applied to a bra gives another bra,

$$\underbrace{\langle 1|}_{\text{bra}}\underbrace{\Big(|2\rangle\langle 3|\Big)}_{\text{ket-bra}} = \underbrace{\langle 1|2\rangle}_{\text{bracket}}\underbrace{\langle 3|}_{\text{bra}}\,. \tag{2.8.14}$$

Splendid, except for the terminology: Rather than of ket-bras, one commonly speaks of *operators*. And then, the unit symbol in (2.8.11) and (2.8.12) is the *identity operator* that leaves kets and bras unchanged,

$$1\big|\,\big\rangle = \big|\,\big\rangle\,, \qquad \big\langle\,\big|1 = \big\langle\,\big|\,, \tag{2.8.15}$$

yet another use of the symbol 1, supplementing the two uses noted after (2.4.25).

Consistent with the rules about adjoints of kets and bras, such as (2.7.8), we have

$$\Big(\big|1\big\rangle\big\langle 2\big|\Big)^\dagger = \big|2\big\rangle\big\langle 1\big| \tag{2.8.16}$$

as the basic rule for adjoints of ket-bras or operators. This is a particular case of the general statement that

$$\begin{aligned}&\text{the adjoint of a product is the product}\\&\text{of the adjoints in reverse order.}\end{aligned} \tag{2.8.17}$$

Here,

$$\Big(\big|1\big\rangle\big\langle 2\big|\Big)^\dagger = \big\langle 2\big|^\dagger\,\big|1\big\rangle^\dagger\,, \tag{2.8.18}$$

and consistency requires that

$$\big\langle 1\big|2\big\rangle = \big\langle 2\big|1\big\rangle^* \tag{2.8.19}$$

holds for any bra $\langle 1|$ and ket $|2\rangle$; see Exercise 13.

What is true for z kets and bras in (2.8.12) must also hold for x kets and bras,

$$\big|{\uparrow_x}\big\rangle\big\langle{\uparrow_x}\big| + \big|{\downarrow_x}\big\rangle\big\langle{\downarrow_x}\big| = 1\,, \tag{2.8.20}$$

which we can verify by translating it into a numerical statement with the aid of the standard rows and columns:

$$\begin{aligned}&\frac{1}{\sqrt2}\begin{pmatrix}1\\1\end{pmatrix}\frac{1}{\sqrt2}\begin{pmatrix}1&1\end{pmatrix} + \frac{1}{\sqrt2}\begin{pmatrix}1\\-1\end{pmatrix}\frac{1}{\sqrt2}\begin{pmatrix}1&-1\end{pmatrix}\\&= \frac12\begin{pmatrix}1&1\\1&1\end{pmatrix} + \frac12\begin{pmatrix}1&-1\\-1&1\end{pmatrix} = \begin{pmatrix}1&0\\0&1\end{pmatrix},\end{aligned} \tag{2.8.21}$$

indeed, and likewise for y kets and bras; see Exercise 14. These statements, the one for x in (2.8.20), that for z in (2.8.12), and the one for y in Exercise 14, are examples of *completeness relations* (or *closure relations*) for the ket and bra bases; the two kinds of bases are paired in the completeness

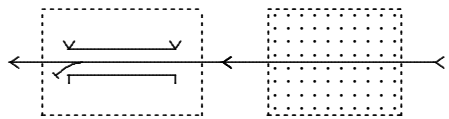
relations so that the completeness of a ket basis relies on the completeness of the partner bra basis and *vice versa*. The completeness is, in fact, a defining property of a basis. For example, two of the unit vectors in (2.7.2), such as $\boldsymbol{e}_x$ and $\boldsymbol{e}_y$, do not make up a basis; they are orthonormal but not complete as there is a third direction in space. So, a completeness relation expresses that we are not missing something. All possibilities are accounted for: "+ in x" *and* "− in x" (or y, or z).

2.9 Pauli operators, Pauli matrices

What about other products of kets and bras such as $|{\uparrow_z}\rangle\langle{\downarrow_z}|$? This operator represents an apparatus that rejects $\uparrow_z$ but accepts $\downarrow_z$ and turns it into $\uparrow_z$,

$$\begin{aligned}\Big(|{\uparrow_z}\rangle\langle{\downarrow_z}|\Big)|{\uparrow_z}\rangle &= |{\uparrow_z}\rangle\underbrace{\langle{\downarrow_z}|{\uparrow_z}\rangle}_{=0} = 0\,,\\ \Big(|{\uparrow_z}\rangle\langle{\downarrow_z}|\Big)|{\downarrow_z}\rangle &= |{\uparrow_z}\rangle\underbrace{\langle{\downarrow_z}|{\downarrow_z}\rangle}_{=1} = |{\uparrow_z}\rangle\,.\end{aligned} \tag{2.9.1}$$

It could be realized by



(2.9.2)

for example. And similarly,

$$|{\uparrow_z}\rangle\langle{\downarrow_z}| + |{\downarrow_z}\rangle\langle{\uparrow_z}| \tag{2.9.3}$$

interchanges $\uparrow_z$ and $\downarrow_z$

$$\begin{aligned}\Big(|{\uparrow_z}\rangle\langle{\downarrow_z}| + |{\downarrow_z}\rangle\langle{\uparrow_z}|\Big)|{\uparrow_z}\rangle &= |{\downarrow_z}\rangle\,,\\ \Big(|{\uparrow_z}\rangle\langle{\downarrow_z}| + |{\downarrow_z}\rangle\langle{\uparrow_z}|\Big)|{\downarrow_z}\rangle &= |{\uparrow_z}\rangle\,,\end{aligned} \tag{2.9.4}$$

essentially the action of the homogeneous magnetic field alone. Let us see what it does to x states,

$$\Big(|{\uparrow_z}\rangle\langle{\downarrow_z}| + |{\downarrow_z}\rangle\langle{\uparrow_z}|\Big)\underbrace{|{\uparrow_x}\rangle}_{=|{\uparrow_z}\rangle\frac{1}{\sqrt{2}}+|{\downarrow_z}\rangle\frac{1}{\sqrt{2}}} = |{\downarrow_z}\rangle\frac{1}{\sqrt{2}} + |{\uparrow_z}\rangle\frac{1}{\sqrt{2}} = |{\uparrow_x}\rangle\,, \tag{2.9.5}$$

that is, $|{\uparrow_x}\rangle$ is left unaltered;

$$\Big(|{\uparrow_z}\rangle\langle{\downarrow_z}| + |{\downarrow_z}\rangle\langle{\uparrow_z}|\Big)\underbrace{|{\downarrow_x}\rangle}_{=|{\uparrow_z}\rangle\frac{1}{\sqrt{2}}-|{\downarrow_z}\rangle\frac{1}{\sqrt{2}}} = |{\downarrow_z}\rangle\frac{1}{\sqrt{2}} - |{\uparrow_z}\rangle\frac{1}{\sqrt{2}} = -|{\downarrow_x}\rangle\,, \tag{2.9.6}$$

that is, $|{\downarrow_x}\rangle$ gets multiplied by -1. Accordingly, it must be true that this operator is also equal to the sum $|{\uparrow_x}\rangle\langle{\uparrow_x}| - |{\downarrow_x}\rangle\langle{\downarrow_x}|$. Indeed, the operator σ_x defined by either one of the expressions

$$\sigma_x = |{\uparrow_z}\rangle\langle{\downarrow_z}| + |{\downarrow_z}\rangle\langle{\uparrow_z}| = |{\uparrow_x}\rangle\langle{\uparrow_x}| - |{\downarrow_x}\rangle\langle{\downarrow_x}| \tag{2.9.7}$$

is uniquely defined irrespective of which expression we regard as the definition. We can verify that in a variety of manners, simplest perhaps by involving once more the standard numerical columns and rows,

$$\begin{aligned}\begin{pmatrix}0&1\\1&0\end{pmatrix} &= \underbrace{\begin{pmatrix}1\\0\end{pmatrix}\begin{pmatrix}0&1\end{pmatrix}}_{=\left(\begin{smallmatrix}0&1\\0&0\end{smallmatrix}\right)} + \underbrace{\begin{pmatrix}0\\1\end{pmatrix}\begin{pmatrix}1&0\end{pmatrix}}_{=\left(\begin{smallmatrix}0&0\\1&0\end{smallmatrix}\right)}\\ &= \underbrace{\frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\end{pmatrix}\frac{1}{\sqrt{2}}\begin{pmatrix}1&1\end{pmatrix}}_{=\frac{1}{2}\left(\begin{smallmatrix}1&1\\1&1\end{smallmatrix}\right)} - \underbrace{\frac{1}{\sqrt{2}}\begin{pmatrix}1\\-1\end{pmatrix}\frac{1}{\sqrt{2}}\begin{pmatrix}1&-1\end{pmatrix}}_{=\frac{1}{2}\left(\begin{smallmatrix}1&-1\\-1&1\end{smallmatrix}\right)}\,.\end{aligned} \tag{2.9.8}$$

The corresponding operators for y and z are

$$\begin{aligned}\sigma_y &= |{\uparrow_y}\rangle\langle{\uparrow_y}| - |{\downarrow_y}\rangle\langle{\downarrow_y}| \mathrel{\widehat{=}} \begin{pmatrix}0&-\mathrm{i}\\ \mathrm{i}&0\end{pmatrix},\\ \sigma_z &= |{\uparrow_z}\rangle\langle{\uparrow_z}| - |{\downarrow_z}\rangle\langle{\downarrow_z}| \mathrel{\widehat{=}} \begin{pmatrix}1&0\\0&-1\end{pmatrix}.\end{aligned} \tag{2.9.9}$$

The three operators σ_x, σ_y, and σ_z are the so-called *Pauli** *operators* and their standard matrix representations

$$\sigma_x \mathrel{\widehat{=}} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y \mathrel{\widehat{=}} \begin{pmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix}, \quad \sigma_z \mathrel{\widehat{=}} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{2.9.10}$$

are known as *Pauli matrices*. The Pauli operators are associated with the three coordinate axes, but since they themselves are arbitrarily oriented, $\sigma_x, \sigma_y, \sigma_z$ must be the cartesian* components of a *Pauli vector operator* $\boldsymbol{\sigma}$,

$$\boldsymbol{\sigma} = \sigma_x \boldsymbol{e}_x + \sigma_y \boldsymbol{e}_y + \sigma_z \boldsymbol{e}_z \,, \tag{2.9.11}$$

a three-dimensional vector with all its typical properties, whose components are operators, however.

By making repeated use of the numerical representation by the Pauli matrices, one easily verifies the fundamental algebraic properties,

$$\begin{aligned} \text{unit square:} \quad & \sigma_x^2 = 1 \,, \qquad \sigma_y^2 = 1 \,, \qquad \sigma_z^2 = 1 \,; \\ \text{cyclic:} \quad & \sigma_x\sigma_y = \mathrm{i}\sigma_z \,, \quad \sigma_y\sigma_z = \mathrm{i}\sigma_x \,, \quad \sigma_z\sigma_x = \mathrm{i}\sigma_y \,; \\ \text{anticyclic:} \quad & \sigma_y\sigma_x = -\mathrm{i}\sigma_z \,, \, \sigma_z\sigma_y = -\mathrm{i}\sigma_x \,, \, \sigma_x\sigma_z = -\mathrm{i}\sigma_y \,. \end{aligned} \tag{2.9.12}$$

A compact statement that comprises all of these as special cases is

$$\boldsymbol{a} \cdot \boldsymbol{\sigma} \;\; \boldsymbol{b} \cdot \boldsymbol{\sigma} = \boldsymbol{a} \cdot \boldsymbol{b} + \mathrm{i}(\boldsymbol{a} \times \boldsymbol{b}) \cdot \boldsymbol{\sigma} \tag{2.9.13}$$

where $\boldsymbol{a}$, $\boldsymbol{b}$ are arbitrary numerical three-dimensional vectors. Letting $\boldsymbol{a}$ and $\boldsymbol{b}$ be any of the three unit vectors for the x, y, and z directions, the nine statements of (2.9.12) are recovered. But now we have a formulation of the algebraic properties that makes no reference to a particular choice of cartesian axes. Some immediate consequences are the subject matter of Exercises 16–18.

2.10 Functions of Pauli operators

An important implication of (2.9.13) is that any function of $\boldsymbol{\sigma}$, however complicated, can be regarded as a linear function of $\boldsymbol{\sigma}$. That is, if $f(\boldsymbol{\sigma})$ is some (operator-valued) function of $\boldsymbol{\sigma}$, then it is always possible to find a number a_0 and a 3-vector $\boldsymbol{a}$ such that

$$f(\boldsymbol{\sigma}) = a_0 + \boldsymbol{a} \cdot \boldsymbol{\sigma} \,. \tag{2.10.1}$$

---

*Wolfgang PAULI (1900–1958) *René DESCARTES (1596–1650)

Strictly speaking, a_0 is multiplied by the unit operator, but we do not make that explicit in the notation as it is obvious from the context. In expressions such as $a_0|\,\rangle = |\,\rangle a_0$, the operator "$a_0$ times the unit operator" multiplies the ket on the left, while the number a_0 multiplies the ket on the right.

A simple example that illustrates how products of Pauli operators are linear functions of $\boldsymbol{\sigma}$ is $\sigma_x\sigma_y = \mathrm{i}\sigma_z$ and the like, and

$$\underbrace{\sigma_x\sigma_y}_{=-\sigma_y\sigma_x}\sigma_x = -\sigma_y\underbrace{\sigma_x\sigma_x}_{=1} = -\sigma_y \tag{2.10.2}$$

is another. Here, we can also argue as indicated by

$$\underbrace{\sigma_x\sigma_y}_{=\mathrm{i}\sigma_z}\sigma_x = \mathrm{i}\underbrace{\sigma_z\sigma_x}_{=\mathrm{i}\sigma_y} = -\sigma_y\,, \tag{2.10.3}$$

with the same outcome. Other examples are found in Exercises 19–23.

Upon combining, for x states say, the completeness relation

$$\big|{\uparrow_x}\big\rangle\big\langle{\uparrow_x}\big| + \big|{\downarrow_x}\big\rangle\big\langle{\downarrow_x}\big| = 1 \tag{2.10.4}$$

with

$$\sigma_x = \big|{\uparrow_x}\big\rangle\big\langle{\uparrow_x}\big| - \big|{\downarrow_x}\big\rangle\big\langle{\downarrow_x}\big|\,, \tag{2.10.5}$$

we find the linear functions of σ_x for "$+$ in x" selection and "$-$ in x" selection,

$$\frac{1+\sigma_x}{2} = \big|{\uparrow_x}\big\rangle\big\langle{\uparrow_x}\big|\,, \qquad \frac{1-\sigma_x}{2} = \big|{\downarrow_x}\big\rangle\big\langle{\downarrow_x}\big|\,. \tag{2.10.6}$$

Indeed, they act correctly as projection operators, or projectors, that pick out the respective components when acting on an arbitrary ket:

$$|\,\rangle = \big|{\uparrow_x}\big\rangle\alpha + \big|{\downarrow_x}\big\rangle\beta \quad \text{with} \quad \alpha = \big\langle{\uparrow_x}\big|\,\rangle\,, \quad \beta = \big\langle{\downarrow_x}\big|\,\rangle\,. \tag{2.10.7}$$

See

$$\begin{aligned}\frac{1\pm\sigma_x}{2}|\,\rangle &= \frac{1\pm\sigma_x}{2}\big|{\uparrow_x}\big\rangle\alpha + \frac{1\pm\sigma_x}{2}\big|{\downarrow_x}\big\rangle\beta \\ &= \big|{\uparrow_x}\big\rangle\frac{1\pm 1}{2}\alpha + \big|{\downarrow_x}\big\rangle\frac{1\mp 1}{2}\beta = \begin{cases}\big|{\uparrow_x}\big\rangle\alpha\,,\\ \big|{\downarrow_x}\big\rangle\beta\,,\end{cases}\end{aligned} \tag{2.10.8}$$

where we use

$$\sigma_x\big|{\uparrow_x}\big\rangle = \big|{\uparrow_x}\big\rangle\,, \qquad \sigma_x\big|{\downarrow_x}\big\rangle = -\big|{\downarrow_x}\big\rangle\,. \tag{2.10.9}$$

Follow up on this with Exercise 24.

2.11 Eigenvalues, eigenkets, and eigenbras

This particular link between σ_x and the x kets $|\uparrow_x\rangle$ and $|\downarrow_x\rangle$, namely that the application of σ_x just multiplies by $+1$ or -1, respectively, has the mathematical structure of an *eigenvector equation*. More generally, we have an eigenket equation of the form

$$\begin{array}{c} A|a\rangle = |a\rangle a \\ \text{operator} \quad \text{eigenket} \quad \text{eigenvalue} \\ \text{of } A \text{ to eigenvalue } a \end{array} \tag{2.11.1}$$

and

$$\begin{array}{c} a\langle a| = \langle a|A \\ \text{eigenvalue} \quad \text{eigenbra} \quad \text{operator} \\ \text{of } A \text{ to eigenvalue } a \end{array} \tag{2.11.2}$$

is the analogous statement about eigenbras, the eigenbra equation. Exercises 25 and 26 deal with basic, yet important aspects.

We convince ourselves that there are no other eigenvalues of σ_x than $+1$ and -1. Because if there were another one, denoted by λ, we would need to find amplitudes α and β such that

$$\sigma_x\Big(|\uparrow_x\rangle\alpha + |\downarrow_x\rangle\beta\Big) = \Big(|\uparrow_x\rangle\alpha + |\downarrow_x\rangle\beta\Big)\lambda\,, \tag{2.11.3}$$

but the left-hand side is equal to $|\uparrow_x\rangle\alpha - |\downarrow_x\rangle\beta$ so that

$$\alpha = \lambda\alpha\,, \qquad \beta = -\lambda\beta \tag{2.11.4}$$

have to hold simultaneously. It follows that either $\lambda = 1$, α arbitrary, $\beta = 0$ or $\lambda = -1$, $\alpha = 0$, β arbitrary, which are the cases we know already. The apparent third possibility λ arbitrary, $\alpha = 0$, $\beta = 0$ does not count because then $|\ \rangle = 0$, which is to say that we do not have an eigenket at all.

How does this argument look if we begin with

$$|\ \rangle = |\uparrow_z\rangle\alpha + |\downarrow_z\rangle\beta\ ? \tag{2.11.5}$$

Then, $\sigma_x|\uparrow_z\rangle = |\downarrow_z\rangle$ and $\sigma_x|\downarrow_z\rangle = |\uparrow_z\rangle$ imply

$$\Big(|\uparrow_z\rangle\alpha + |\downarrow_z\rangle\beta\Big)\lambda \underset{\substack{\uparrow\\ \text{(want)}}}{=} \sigma_x\Big(|\uparrow_z\rangle\alpha + |\downarrow_z\rangle\beta\Big) = |\downarrow_z\rangle\alpha + |\uparrow_z\rangle\beta\,, \tag{2.11.6}$$

that is,

$$\alpha = \lambda\beta\,, \qquad \beta = \lambda\alpha \tag{2.11.7}$$

or

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \lambda\begin{pmatrix} \alpha \\ \beta \end{pmatrix} \tag{2.11.8}$$

or

$$\begin{pmatrix} -\lambda & 1 \\ 1 & -\lambda \end{pmatrix}\begin{pmatrix} \alpha \\ \beta \end{pmatrix} = 0\,. \tag{2.11.9}$$

The latter is, of course, just the standard numerical version of

$$(\sigma_x - \lambda)|\ \rangle = 0\,. \tag{2.11.10}$$

Now, since $|\ \rangle = 0$, that is, $\alpha = \beta = 0$, is not an option, such a set of coupled linear equations has a solution only if the determinant of the matrix vanishes,

$$0 = \det\left(\begin{pmatrix} -\lambda & 1 \\ 1 & -\lambda \end{pmatrix}\right) = \lambda^2 - 1\,. \tag{2.11.11}$$

This implies first that $\lambda = 1$ or $\lambda = -1$ and then $\alpha = \beta$ or $\alpha = -\beta$, respectively, so that

$$\begin{aligned} |\lambda = 1\rangle &\propto |\uparrow_z\rangle + |\downarrow_z\rangle\,, \\ |\lambda = -1\rangle &\propto |\uparrow_z\rangle - |\downarrow_z\rangle\,, \end{aligned} \tag{2.11.12}$$

with the proportionality factors determined by some convention. We usually insist on kets of unit length so that

$$|\lambda = \pm 1\rangle = \frac{1}{\sqrt{2}}\Big(|\uparrow_z\rangle \pm |\downarrow_z\rangle\Big) = \begin{cases} |\uparrow_x\rangle\,, \\ |\downarrow_x\rangle\,, \end{cases} \tag{2.11.13}$$

hardly a surprising result. By the same token, we have

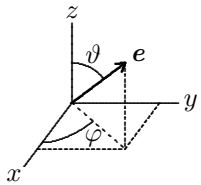
$$\sigma_z|\uparrow_z\rangle = |\uparrow_z\rangle\,, \quad \sigma_z|\downarrow_z\rangle = -|\downarrow_z\rangle \tag{2.11.14}$$

and

$$\sigma_y\big|{\uparrow_y}\big\rangle = \big|{\uparrow_y}\big\rangle\,, \quad \sigma_y\big|{\downarrow_y}\big\rangle = -\big|{\downarrow_y}\big\rangle\,, \tag{2.11.15}$$

and we expect that for *any* component $\boldsymbol{e}\cdot\boldsymbol{\sigma}$ of Pauli's vector operator $\boldsymbol{\sigma}$, there are analogous eigenkets to the same eigenvalues $+1$ and -1.

It is worth checking this in detail, for which purpose we parameterize the *unit vector* $\boldsymbol{e}$ in spherical coordinates,



$$\boldsymbol{e} \mathrel{\widehat{=}} \begin{pmatrix} \sin(\vartheta)\cos(\varphi) \\ \sin(\vartheta)\sin(\varphi) \\ \cos(\vartheta) \end{pmatrix}. \tag{2.11.16}$$

Then,

$$\begin{aligned} \boldsymbol{e}\cdot\boldsymbol{\sigma} &= \sigma_x \sin(\vartheta)\cos(\varphi) + \sigma_y \sin(\vartheta)\sin(\varphi) + \sigma_z \cos(\vartheta) \\ &\mathrel{\widehat{=}} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \sin(\vartheta)\cos(\varphi) + \begin{pmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix} \sin(\vartheta)\sin(\varphi) + \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \cos(\vartheta) \\ &= \begin{pmatrix} \cos(\vartheta) & \mathrm{e}^{-\mathrm{i}\varphi}\sin(\vartheta) \\ \mathrm{e}^{\mathrm{i}\varphi}\sin(\vartheta) & -\cos(\vartheta) \end{pmatrix} \end{aligned} \tag{2.11.17}$$

utilizes the standard numerical representations of σ_x, σ_y, and σ_z once more, and the eigenvector equation

$$\big(\boldsymbol{e}\cdot\boldsymbol{\sigma} - \lambda\big)\big|\;\big\rangle = 0 \tag{2.11.18}$$

appears as

$$\begin{pmatrix} \cos(\vartheta) - \lambda & \mathrm{e}^{-\mathrm{i}\varphi}\sin(\vartheta) \\ \mathrm{e}^{\mathrm{i}\varphi}\sin(\vartheta) & -\cos(\vartheta) - \lambda \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = 0\,. \tag{2.11.19}$$

The determinant of this 2×2 matrix vanishes when λ is an eigenvalue,

$$0 = \big(\cos(\vartheta) - \lambda\big)\big(-\cos(\vartheta) - \lambda\big) - \big(\mathrm{e}^{-\mathrm{i}\varphi}\sin(\vartheta)\big)\big(\mathrm{e}^{\mathrm{i}\varphi}\sin(\vartheta)\big) = \lambda^2 - 1\,, \tag{2.11.20}$$

so that $\lambda = +1$ and $\lambda = -1$ are the eigenvalues, indeed. To find the respective eigenkets, we solve for the amplitudes α, β after choosing one of the eigenvalues. Let us do it for $\lambda = 1$,

$$\begin{pmatrix} \cos(\vartheta) - 1 & \mathrm{e}^{-\mathrm{i}\varphi}\sin(\vartheta) \\ \mathrm{e}^{\mathrm{i}\varphi}\sin(\vartheta) & -\cos(\vartheta) - 1 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = 0\,, \tag{2.11.21}$$

or, after introducing $\frac{1}{2}\vartheta$,

$$-2\begin{pmatrix} \sin\left(\frac{1}{2}\vartheta\right)^2 & -\mathrm{e}^{-\mathrm{i}\varphi}\sin\left(\frac{1}{2}\vartheta\right)\cos\left(\frac{1}{2}\vartheta\right) \\ -\mathrm{e}^{\mathrm{i}\varphi}\sin\left(\frac{1}{2}\vartheta\right)\cos\left(\frac{1}{2}\vartheta\right) & \cos\left(\frac{1}{2}\vartheta\right)^2 \end{pmatrix}\begin{pmatrix}\alpha\\ \beta\end{pmatrix} = 0\,, \tag{2.11.22}$$

which can also be written as

$$-2\begin{pmatrix} -\mathrm{e}^{-\mathrm{i}\varphi}\sin\left(\frac{1}{2}\vartheta\right) \\ \cos\left(\frac{1}{2}\vartheta\right)\end{pmatrix}\begin{pmatrix} -\mathrm{e}^{\mathrm{i}\varphi}\sin\left(\frac{1}{2}\vartheta\right) & \cos\left(\frac{1}{2}\vartheta\right)\end{pmatrix}\begin{pmatrix}\alpha\\ \beta\end{pmatrix} = 0\,. \tag{2.11.23}$$

Now, the column on the left is never $\begin{pmatrix}0\\0\end{pmatrix}$ and, therefore, the product of the central row with the column on the right must vanish, which requires

$$\alpha\,\mathrm{e}^{\mathrm{i}\varphi}\sin\left(\frac{1}{2}\vartheta\right) = \beta\cos\left(\frac{1}{2}\vartheta\right) \tag{2.11.24}$$

and implies

$$\begin{pmatrix}\alpha\\ \beta\end{pmatrix} = \begin{pmatrix}\cos\left(\frac{1}{2}\vartheta\right)\\ \mathrm{e}^{\mathrm{i}\varphi}\sin\left(\frac{1}{2}\vartheta\right)\end{pmatrix}, \tag{2.11.25}$$

up to an over-all phase factor, which we choose to equal 1. The eigenket to $\lambda = +1$ is thus

$$\left|\uparrow_{\boldsymbol{e}}\right\rangle = \left|\uparrow_z\right\rangle\cos\left(\frac{1}{2}\vartheta\right) + \left|\downarrow_z\right\rangle\mathrm{e}^{\mathrm{i}\varphi}\sin\left(\frac{1}{2}\vartheta\right). \tag{2.11.26}$$

Rather than repeating the argument for $\lambda = -1$, we note that we have just looked at

$$(\boldsymbol{e}\cdot\boldsymbol{\sigma} - 1)\left|\uparrow_{\boldsymbol{e}}\right\rangle = 0\,, \tag{2.11.27}$$

where

$$\boldsymbol{e}\cdot\boldsymbol{\sigma} - 1 = -2\,\frac{1-\boldsymbol{e}\cdot\boldsymbol{\sigma}}{2} = -2\left|\downarrow_{\boldsymbol{e}}\right\rangle\left\langle\downarrow_{\boldsymbol{e}}\right| \tag{2.11.28}$$

in analogy with $|\downarrow_x\rangle\langle\downarrow_x| = \frac{1}{2}(1-\sigma_x)$, for example.

Accordingly, we have

$$-2\left|\downarrow_{\boldsymbol{e}}\right\rangle\left\langle\downarrow_{\boldsymbol{e}}\right| \mathrel{\widehat{=}} -2\begin{pmatrix} -\mathrm{e}^{-\mathrm{i}\varphi}\sin\left(\frac{1}{2}\vartheta\right) \\ \cos\left(\frac{1}{2}\vartheta\right)\end{pmatrix}\begin{pmatrix} -\mathrm{e}^{\mathrm{i}\varphi}\sin\left(\frac{1}{2}\vartheta\right) & \cos\left(\frac{1}{2}\vartheta\right)\end{pmatrix} \tag{2.11.29}$$

and read off that

$$\left|\downarrow_{\boldsymbol{e}}\right\rangle \mathrel{\widehat{=}} \begin{pmatrix} -\mathrm{e}^{-\mathrm{i}\varphi}\sin\left(\frac{1}{2}\vartheta\right) \\ \cos\left(\frac{1}{2}\vartheta\right)\end{pmatrix} \tag{2.11.30}$$

or

$$\big|{\downarrow_{\boldsymbol{e}}}\big\rangle = \big|{\uparrow_z}\big\rangle\Big(-\,\mathrm{e}^{-\mathrm{i}\varphi}\sin\Big(\frac{1}{2}\vartheta\Big)\Big) + \big|{\downarrow_z}\big\rangle\cos\Big(\frac{1}{2}\vartheta\Big)\,. \tag{2.11.31}$$

It is a matter of inspection to verify that

$$\boldsymbol{e}\cdot\boldsymbol{\sigma}\,\big|{\uparrow_{\boldsymbol{e}}}\big\rangle = \big|{\uparrow_{\boldsymbol{e}}}\big\rangle \quad\text{and}\quad \boldsymbol{e}\cdot\boldsymbol{\sigma}\,\big|{\downarrow_{\boldsymbol{e}}}\big\rangle = -\big|{\downarrow_{\boldsymbol{e}}}\big\rangle \tag{2.11.32}$$

and that the orthogonality relation

$$\big\langle{\uparrow_{\boldsymbol{e}}}\big|{\downarrow_{\boldsymbol{e}}}\big\rangle = 0 \tag{2.11.33}$$

holds. Also, we have built in the normalization,

$$\big\langle{\uparrow_{\boldsymbol{e}}}\big|{\uparrow_{\boldsymbol{e}}}\big\rangle = 1\,, \qquad \big\langle{\downarrow_{\boldsymbol{e}}}\big|{\downarrow_{\boldsymbol{e}}}\big\rangle = 1\,, \tag{2.11.34}$$

and the generalization of the completeness relations (2.8.12) and (2.8.20) is

$$\big|{\uparrow_{\boldsymbol{e}}}\big\rangle\big\langle{\uparrow_{\boldsymbol{e}}}\big| + \big|{\downarrow_{\boldsymbol{e}}}\big\rangle\big\langle{\downarrow_{\boldsymbol{e}}}\big| = 1\,. \tag{2.11.35}$$

It follows that the pair $|{\uparrow_{\boldsymbol{e}}}\rangle, |{\downarrow_{\boldsymbol{e}}}\rangle$ is another basis for the kets and $\langle{\uparrow_{\boldsymbol{e}}}|, \langle{\downarrow_{\boldsymbol{e}}}|$ is another basis for the bras.

2.12 Wave–particle duality

All this rather simple mathematics may not be so impressive, but there really is a fundamental physical insight in it. There is a *continuum* of $\boldsymbol{e}\cdot\boldsymbol{\sigma}$ operators, each referring to magnetic properties of the silver atoms associated with the specified direction $\boldsymbol{e}$ in space, but for each direction, the values are *discrete*, invariably $+1$ and -1.

This union of the discrete and the continuous is at the heart of quantum mechanics. Discrete features are associated with particle-like objects in classical physics (planets in orbit, falling stones, swinging pendulums, ...), whereas continuous features are associated with wave phenomena (water waves, sound waves, radio waves, ...). The manifestation of both in the same class of phenomena triggered the historical debate about *wave–particle duality* that commenced with Einstein's paper of 1905 on the photoelectric effect and still continues today.

This example of continuous directional options with discrete alternatives for each of them is the paradigmatic illustration of "waves and particles in peaceful coexistence" or the discrete-and-continuous. And please do not fail to note how this came about as a natural, unavoidable consequence of the experimental fact that atoms are deflected up *or* down in a Stern–Gerlach experiment, with no other option available to them.

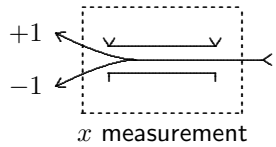
2.13 Expectation value

We observed above that the Pauli operators are of the form

$$\left.\begin{matrix}\sigma_x\\ \sigma_y\\ \sigma_z\end{matrix}\right\} = \left\{\begin{matrix}|{\uparrow_x}\rangle\langle{\uparrow_x}| - |{\downarrow_x}\rangle\langle{\downarrow_x}|\\ |{\uparrow_y}\rangle\langle{\uparrow_y}| - |{\downarrow_y}\rangle\langle{\downarrow_y}|\\ |{\uparrow_z}\rangle\langle{\uparrow_z}| - |{\downarrow_z}\rangle\langle{\downarrow_z}|\end{matrix}\right\}$$

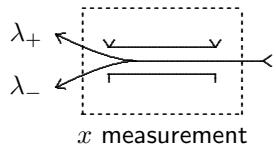
$$= (+1)(+\text{selector}) + (-1)(-\text{selector}) \tag{2.13.1}$$

and noted that the numbers $+1$ and -1 appearing here are the eigenvalues of these operators. We have, in effect, introduced a mathematical symbol that represents the physical property measured by the respective Stern–Gerlach apparatus, with the understanding that we assign $+1$ to "deflected up" and -1 to "deflected down,"



(2.13.2)

While this is the conventional, and perhaps most natural, assignment of numbers that characterize the measurement result, it is by no means unique. Just as well we could have chosen any other two numbers, real or complex. If we denote them by λ_+ and λ_-, we have this situation:



(2.13.3)

and the corresponding symbol is

$$\Lambda = |{\uparrow_x}\rangle\lambda_+\langle{\uparrow_x}| + |{\downarrow_x}\rangle\lambda_-\langle{\downarrow_x}| \tag{2.13.4}$$

or, if we recall that $\frac{1}{2}(1 \pm \sigma_x)$ are the projectors in question,

$$\Lambda = \lambda_+\frac{1+\sigma_x}{2} + \lambda_-\frac{1-\sigma_x}{2} = \frac{1}{2}(\lambda_+ + \lambda_-) + \frac{1}{2}(\lambda_+ - \lambda_-)\sigma_x\,. \tag{2.13.5}$$

The kets $|\uparrow_x\rangle$ and $|\downarrow_x\rangle$ remain eigenkets, but the eigenvalues are now λ_+ and λ_-,

$$\Lambda\big|\uparrow_x\big\rangle = \big|\uparrow_x\big\rangle\lambda_+\,, \qquad \Lambda\big|\downarrow_x\big\rangle = \big|\downarrow_x\big\rangle\lambda_-\,. \tag{2.13.6}$$

There is nothing in the physics of the situation that would dictate one particular choice for λ_+ and λ_-, and so all Λ operators constructed in this way refer to the magnetic property of the silver atoms measured by a Stern–Gerlach apparatus for the x direction. Put differently, an apparatus that measures σ_x measures at the same time all functions of σ_x — recall here that the most general function is a constant added to a multiple of σ_x, a linear function of σ_x.

Now, having opted for a particular λ_+ and λ_- pair, the probability of getting outcome λ_+ is just the probability for up-deflection,

$$\text{prob}(\lambda_+) = \big|\big\langle\uparrow_x\big|\,\big\rangle\big|^2\,, \tag{2.13.7}$$

and likewise

$$\text{prob}(\lambda_-) = \big|\big\langle\downarrow_x\big|\,\big\rangle\big|^2 \tag{2.13.8}$$

for λ_- and down-deflection.

An individual atom will either be recorded with value λ_+ or value λ_-, and these probabilities state the relative frequencies with which the values occur if very many atoms are measured. The average value,

$$\lambda_+\,\text{prob}(\lambda_+) + \lambda_-\,\text{prob}(\lambda_-)\,, \tag{2.13.9}$$

is the weighted sum of the two measurement results. Introducing the probabilities of (2.13.7) and (2.13.8), we have

$$\begin{aligned}\lambda_+\big|\big\langle\uparrow_x\big|\,\big\rangle\big|^2 + \lambda_-\big|\big\langle\downarrow_x\big|\,\big\rangle\big|^2 &= \big\langle\,\big|\uparrow_x\big\rangle\lambda_+\big\langle\uparrow_x\big|\,\big\rangle + \big\langle\,\big|\downarrow_x\big\rangle\lambda_-\big\langle\downarrow_x\big|\,\big\rangle \\ &= \big\langle\,\Big|\Big(\big|\uparrow_x\big\rangle\lambda_+\big\langle\uparrow_x\big| + \big|\downarrow_x\big\rangle\lambda_-\big\langle\downarrow_x\big|\Big)\big|\,\big\rangle \\ &= \big\langle\,\big|\Lambda\big|\,\big\rangle\,. \end{aligned} \tag{2.13.10}$$

So, the eigenvalues of Λ are the measurement results, and their statistical mean value is the number that we obtain by sandwiching Λ by the bra $\langle\,|$ and the ket $|\,\rangle$ that describe the atoms that are measured. A common shorthand notation is $\langle\Lambda\rangle$ for such an average,

$$\langle\Lambda\rangle = \lambda_+\text{prob}(\lambda_+) + \lambda_-\text{prob}(\lambda_-) = \big\langle\,\big|\Lambda\big|\,\big\rangle\,, \tag{2.13.11}$$

and standard terminology is to call it the *mean value of* Λ or the *expectation value of* Λ. The latter is very widespread, almost universal terminology. The statistics terminology *expected value* is less common in physics; we shall not use it, also because $\langle\Lambda\rangle$ is not a "value that we expect." Rather, we expect to get either λ_+ or λ_-, that is, one of the eigenvalues of Λ, in each measurement act and the statistical average of those numbers is the expectation value of Λ.

Probabilities are expectation values as well:

$$\begin{aligned}\mathrm{prob}(\lambda_+) &= \big|\langle\uparrow_x|\ \rangle\big|^2 = \langle\ |\Big(|\uparrow_x\rangle\langle\uparrow_x|\Big)|\ \rangle = \langle\ |\frac{1+\sigma_x}{2}|\ \rangle \\ &= \left\langle\frac{1+\sigma_x}{2}\right\rangle\,, \end{aligned} \tag{2.13.12}$$

and likewise

$$\mathrm{prob}(\lambda_-) = \left\langle\frac{1-\sigma_x}{2}\right\rangle\,. \tag{2.13.13}$$

With this in mind, it should be clear that *all* statistical measurement results are expectation values of the appropriate operators. Probabilities, in particular, are the expectation values of the operators that select the outcome, such as $\frac{1}{2}(1\pm\sigma_x)$ in our example.

2.14 Trace

To proceed further, we need to introduce a new mathematical operation, called the *trace*. It assigns a number to each operator, the basic relation being

$$\mathrm{tr}\Big(|1\rangle\langle 2|\Big) = \langle 2|1\rangle\,, \tag{2.14.1}$$

pronounced: the trace of the ket-bra $|1\rangle\langle 2|$ is the bra-ket $\langle 2|1\rangle$. Since it is a single number, this is not a reversible mathematical operation. Knowing the operator, you can work out the trace, but from the knowledge of the trace alone, you cannot infer the operator uniquely. The trace is a *linear operation*,

$$\begin{aligned}\mathrm{tr}(X+Y) &= \mathrm{tr}(X)+\mathrm{tr}(Y)\,,\\ \mathrm{tr}(\lambda X) &= \lambda\,\mathrm{tr}(X)\,,\end{aligned} \tag{2.14.2}$$

for all operators X, Y and all complex numbers λ. This linearity is necessary for consistency with the linear structure of the ket and bra spaces, as illustrated by

$$|1\rangle = |3\rangle\alpha + |4\rangle\beta \tag{2.14.3}$$

and

$$\begin{aligned} \mathrm{tr}\Big(|1\rangle\langle 2|\Big) &= \langle 2|1\rangle = \langle 2|3\rangle\alpha + \langle 2|4\rangle\beta \\ &= \alpha\,\mathrm{tr}\Big(|3\rangle\langle 2|\Big) + \beta\,\mathrm{tr}\Big(|4\rangle\langle 2|\Big) \end{aligned} \tag{2.14.4}$$

and the like.

For an arbitrary operator X, we can use *some* decomposition of the identity,

$$|\uparrow\rangle\langle\uparrow| + |\downarrow\rangle\langle\downarrow| = 1\,, \tag{2.14.5}$$

to establish

$$\begin{aligned} \mathrm{tr}(X) &= \mathrm{tr}\bigg(\Big(|\uparrow\rangle\langle\uparrow| + |\downarrow\rangle\langle\downarrow|\Big)X\bigg) \\ &= \mathrm{tr}\bigg(|\uparrow\rangle\Big(\langle\uparrow|X\Big) + |\downarrow\rangle\Big(\langle\downarrow|X\Big)\bigg) \\ &= \langle\uparrow|X|\uparrow\rangle + \langle\downarrow|X|\downarrow\rangle\,, \end{aligned} \tag{2.14.6}$$

which says the following: If we represent X by a 2×2 matrix in accordance with the usual procedure,

$$\begin{aligned} X = 1X1 &= \Big(|\uparrow\rangle\langle\uparrow| + |\downarrow\rangle\langle\downarrow|\Big)X\Big(|\uparrow\rangle\langle\uparrow| + |\downarrow\rangle\langle\downarrow|\Big) \\ &= \underbrace{\Big(\,|\uparrow\rangle\;\;|\downarrow\rangle\,\Big)}_{\text{row of kets}}\underbrace{\begin{pmatrix}\langle\uparrow|X|\uparrow\rangle & \langle\uparrow|X|\downarrow\rangle\\ \langle\downarrow|X|\uparrow\rangle & \langle\downarrow|X|\downarrow\rangle\end{pmatrix}}_{2\times 2\text{ matrix}}\underbrace{\begin{pmatrix}\langle\uparrow|\\ \langle\downarrow|\end{pmatrix}}_{\text{column of bras}} \end{aligned} \tag{2.14.7}$$

or

$$X \mathrel{\widehat{=}} \begin{pmatrix}\langle\uparrow|X|\uparrow\rangle & \langle\uparrow|X|\downarrow\rangle\\ \langle\downarrow|X|\uparrow\rangle & \langle\downarrow|X|\downarrow\rangle\end{pmatrix} \tag{2.14.8}$$

for short, then $\mathrm{tr}(X)$ is simply the sum of the diagonal matrix elements. Note that this statement is true irrespective of whether the reference states are up/down in x, or y, or z, or any other direction; see Exercise 32 for an example.

In addition to its linearity, the trace operation has another important property, namely that

$$\mathrm{tr}(XY) = \mathrm{tr}(YX) \tag{2.14.9}$$

for any two operators X and Y. In view of the linearity, it suffices to consider $X = |1\rangle\langle 2|$ and $Y = |3\rangle\langle 4|$,

$$\begin{aligned}\mathrm{tr}(XY) &= \mathrm{tr}\Big(|1\rangle\langle 2|3\rangle\langle 4|\Big)\\ &= \langle 2|3\rangle\,\mathrm{tr}\Big(|1\rangle\langle 4|\Big) = \langle 2|3\rangle\langle 4|1\rangle\\ &= \mathrm{tr}\Big(|3\rangle\langle 2|\Big)\,\langle 4|1\rangle\\ &= \mathrm{tr}\Big(|3\rangle\langle 4|1\rangle\langle 2|\Big) = \mathrm{tr}(YX),\end{aligned} \tag{2.14.10}$$

indeed.

As an immediate consequence, we note the *cyclic property of the trace*,

$$\mathrm{tr}(XYZ) = \mathrm{tr}(YZX) = \mathrm{tr}(ZXY) \tag{2.14.11}$$

for any three operators X, Y, and Z, and analogous statements about the cyclic permutation of factors hold for four, five, or more factors.

Now, having added the trace operation to our toolbox, let us reconsider the expectation value of operator Λ,

$$\langle\Lambda\rangle = \langle\ |\Lambda|\ \rangle = \langle\ |\Big(\Lambda|\ \rangle\Big) = \mathrm{tr}\Big(\Big(\Lambda|\ \rangle\Big)\langle\ |\Big)\,, \tag{2.14.12}$$

that is,

$$\langle\Lambda\rangle = \mathrm{tr}\Big(\Lambda\,|\ \rangle\langle\ |\Big) = \mathrm{tr}\Big(|\ \rangle\langle\ |\,\Lambda\Big)\,, \tag{2.14.13}$$

where we succeeded in separating the information about the atoms measured, contained in $|\ \rangle\langle\ |$, from the information about the measured property, symbolized by Λ.

2.15 Statistical operator, Born rule

To go beyond mere rewriting of previously obtained expressions, we get to something new by asking the following question: How do we describe atoms prepared by a mixing procedure, such as 50% as "+ in x" and 50% as "+ in z"? We answer the question by looking at the resulting expectation value of Λ,

$$\begin{aligned}
\text{atoms “+ in } x\text{”} : \ \langle\Lambda\rangle &= \mathrm{tr}\Big(\Lambda|{\uparrow_x}\rangle\langle{\uparrow_x}|\Big)\,, \\
\text{atoms “+ in } z\text{”} : \ \langle\Lambda\rangle &= \mathrm{tr}\Big(\Lambda|{\uparrow_z}\rangle\langle{\uparrow_z}|\Big)\,, \\
50\%\text{ of each kind} : \ \langle\Lambda\rangle &= \frac{1}{2}\mathrm{tr}\Big(\Lambda|{\uparrow_x}\rangle\langle{\uparrow_x}|\Big) + \frac{1}{2}\mathrm{tr}\Big(\Lambda|{\uparrow_z}\rangle\langle{\uparrow_z}|\Big) \\
&= \mathrm{tr}\bigg(\Lambda\Big(\frac{1}{2}|{\uparrow_x}\rangle\langle{\uparrow_x}| + \frac{1}{2}|{\uparrow_z}\rangle\langle{\uparrow_z}|\Big)\bigg) \qquad (2.15.1)
\end{aligned}$$

or

$$\langle\Lambda\rangle = \mathrm{tr}(\Lambda\rho) \qquad (2.15.2)$$

with

$$\rho = \frac{1}{2}\Big(|{\uparrow_x}\rangle\langle{\uparrow_x}| + |{\uparrow_z}\rangle\langle{\uparrow_z}|\Big). \qquad (2.15.3)$$

This new mathematical object ρ, which summarizes all we know about the preparation of the atoms, is called *statistical operator* or *state operator* (sometimes “probability operator”) and, somewhat misleadingly but rather commonly, *density matrix*. Our preferred terminology is *statistical operator* or simply *state*.

The statement in (2.15.2), and earlier in (2.14.13), with the trace of the product of the statistical operator and the property operator — here, ρ or $|\,\rangle\langle\,|$ and Λ, respectively — is the *Born* rule*. It is one of the links, arguably the most important one, that connect the mathematical formalism of quantum mechanics with the phenomena it deals with.

As all operators, also the statistical operator in (2.15.3) is a function of Pauli’s vector operator $\boldsymbol{\sigma}$, a linear function in fact. For the example above, we have

$$\rho = \frac{1}{2}\frac{1+\sigma_x}{2} + \frac{1}{2}\frac{1+\sigma_z}{2} = \frac{1}{2}\bigg(1 + \frac{1}{2}\sigma_x + \frac{1}{2}\sigma_z\bigg), \qquad (2.15.4)$$

and in general we get

$$\rho = a_0 + \boldsymbol{a}\cdot\boldsymbol{\sigma} \qquad (2.15.5)$$

with a numerical vector $\boldsymbol{a}$ and a number a_0 (a multiple of the identity). The probabilities associated with an arbitrary direction are the expectation values of $|{\uparrow}\rangle\langle{\uparrow}|$ and $|{\downarrow}\rangle\langle{\downarrow}|$,

*Max BORN (1882–1970)

$$\begin{aligned}\text{prob}(\text{“up”}) &= \text{tr}\Big(|{\uparrow}\rangle\langle{\uparrow}|\rho\Big) = \langle{\uparrow}|\rho|{\uparrow}\rangle\,,\\ \text{prob}(\text{“down”}) &= \text{tr}\Big(|{\downarrow}\rangle\langle{\downarrow}|\rho\Big) = \langle{\downarrow}|\rho|{\downarrow}\rangle\,,\end{aligned} \tag{2.15.6}$$

and their unit sum

$$\begin{aligned}1 &= \text{prob}(\text{“up”}) + \text{prob}(\text{“down”})\\ &= \text{tr}\Big(\underbrace{\Big(|{\uparrow}\rangle\langle{\uparrow}| + |{\downarrow}\rangle\langle{\downarrow}|\Big)}_{=1}\rho\Big) = \text{tr}(\rho)\,,\end{aligned} \tag{2.15.7}$$

or

$$1 = \langle{\uparrow}|\rho|{\uparrow}\rangle + \langle{\downarrow}|\rho|{\downarrow}\rangle = \text{tr}(\rho)\,, \tag{2.15.8}$$

states

$$\text{tr}(\rho) = 1 \tag{2.15.9}$$

as the condition for unit total probability. Now, the Pauli operators $\sigma_x, \sigma_y, \sigma_z$ are traceless,

$$\text{tr}(\sigma_x) = 0\,,\quad \text{tr}(\sigma_y) = 0\,,\quad \text{tr}(\sigma_z) = 0\,,\quad \text{tr}(\boldsymbol{\sigma}) = 0\,, \tag{2.15.10}$$

which is easily verified by a look at the diagonal sum of their standard matrix representations in (2.9.10), and the trace of the identity operator is 2,

$$\text{tr}(1) = 2\,, \tag{2.15.11}$$

see Exercise 31. Accordingly, we obtain

$$\text{tr}(\rho) = 2a_0 = 1 \quad \text{or} \quad a_0 = \frac{1}{2}\,. \tag{2.15.12}$$

Further consider

$$\begin{aligned}\langle\sigma_x\rangle &= \text{tr}(\sigma_x\rho) = \text{tr}\big(\sigma_x(a_0 + \boldsymbol{a}\cdot\boldsymbol{\sigma})\big)\\ &= a_0\underbrace{\text{tr}(\sigma_x)}_{=0} + \text{tr}\big(a_x + \underbrace{a_y\text{i}\sigma_z}_{\to 0} + \underbrace{a_z(-\text{i}\sigma_y)}_{\to 0}\big)\end{aligned} \tag{2.15.13}$$

so that $\langle\sigma_x\rangle = 2a_x$ and likewise $\langle\sigma_y\rangle = 2a_y$, $\langle\sigma_z\rangle = 2a_z$, compactly summarized in

$$\langle\boldsymbol{\sigma}\rangle = 2\boldsymbol{a}\,. \tag{2.15.14}$$

Thus, the general statistical operator is of the form

$$\rho = \frac{1}{2}(1 + \boldsymbol{s}\cdot\boldsymbol{\sigma}) \quad \text{with} \quad \boldsymbol{s} = \langle\boldsymbol{\sigma}\rangle = \mathrm{tr}(\boldsymbol{\sigma}\rho)\,, \tag{2.15.15}$$

where we meet the *Bloch** *vector* $\boldsymbol{s}$.

What then are the probabilities for "up" and "down" in an arbitrary direction specified by unit vector $\boldsymbol{e}$? Since the respective projectors are

$$|{\uparrow}\rangle\langle{\uparrow}| = \frac{1}{2}(1 + \boldsymbol{e}\cdot\boldsymbol{\sigma}) \quad \text{and} \quad |{\downarrow}\rangle\langle{\downarrow}| = \frac{1}{2}(1 - \boldsymbol{e}\cdot\boldsymbol{\sigma})\,, \tag{2.15.16}$$

we have

$$\left.\begin{matrix}\mathrm{prob}(\text{“up”})\\ \mathrm{prob}(\text{“down”})\end{matrix}\right\} = \mathrm{tr}\left(\frac{1}{2}(1 \pm \boldsymbol{e}\cdot\boldsymbol{\sigma})\rho\right) = \mathrm{tr}\left(\frac{1}{2}(1 \pm \boldsymbol{e}\cdot\boldsymbol{\sigma})\frac{1}{2}(1 + \boldsymbol{s}\cdot\boldsymbol{\sigma})\right)$$
$$= \mathrm{tr}\left(\frac{1}{4} \pm \frac{1}{4}\boldsymbol{e}\cdot\boldsymbol{\sigma} + \frac{1}{4}\boldsymbol{s}\cdot\boldsymbol{\sigma} \pm \frac{1}{4}\boldsymbol{e}\cdot\boldsymbol{\sigma}\ \boldsymbol{s}\cdot\boldsymbol{\sigma}\right), \tag{2.15.17}$$

that is,

$$\left.\begin{matrix}\mathrm{prob}(\text{“up”})\\ \mathrm{prob}(\text{“down”})\end{matrix}\right\} = \frac{1}{2} \pm \frac{1}{4}\mathrm{tr}(\boldsymbol{e}\cdot\boldsymbol{\sigma}\ \boldsymbol{s}\cdot\boldsymbol{\sigma})\,. \tag{2.15.18}$$

We evaluate the remaining trace by first recalling that

$$\boldsymbol{e}\cdot\boldsymbol{\sigma}\ \ \boldsymbol{s}\cdot\boldsymbol{\sigma} = \boldsymbol{e}\cdot\boldsymbol{s} + \mathrm{i}(\boldsymbol{e}\times\boldsymbol{s})\cdot\boldsymbol{\sigma}\,, \tag{2.15.19}$$

see (2.9.13), and noting once more that $\mathrm{tr}(1) = 2$, $\mathrm{tr}(\boldsymbol{\sigma}) = 0$, with the consequence

$$\mathrm{tr}(\boldsymbol{e}\cdot\boldsymbol{\sigma}\ \boldsymbol{s}\cdot\boldsymbol{\sigma}) = 2\boldsymbol{e}\cdot\boldsymbol{s} \tag{2.15.20}$$

and then

$$\left.\begin{matrix}\mathrm{prob}(\text{“up”})\\ \mathrm{prob}(\text{“down”})\end{matrix}\right\} = \frac{1}{2} \pm \frac{1}{2}\boldsymbol{e}\cdot\boldsymbol{s} = \frac{1}{2}(1 \pm \boldsymbol{e}\cdot\boldsymbol{s})\,. \tag{2.15.21}$$

Both probabilities are numbers in the range $0\cdots1$ so that

$$-1 \le \boldsymbol{e}\cdot\boldsymbol{s} \le 1 \tag{2.15.22}$$

for all unit vectors $\boldsymbol{e}$ and all permissible vectors $\boldsymbol{s}$. It follows that the length of $\boldsymbol{s}$ cannot exceed unity,

$$|\boldsymbol{s}| = \sqrt{\boldsymbol{s}\cdot\boldsymbol{s}} \le 1\,. \tag{2.15.23}$$

---

*Felix BLOCH (1905–1985)

2.16 Mixtures and blends

For the example of (2.15.4), we have $\boldsymbol{s} = \frac{1}{2}(\boldsymbol{e}_1 + \boldsymbol{e}_3) \mathrel{\widehat{=}} \frac{1}{2}\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$. We can also write this ρ as

$$\rho = \frac{3+\sqrt{3}}{6}\,\frac{1+\frac{1}{2}\big(\sigma_x+\sqrt{3}\sigma_z\big)}{2} + \frac{3-\sqrt{3}}{6}\,\frac{1+\frac{1}{2}\big(\sigma_x-\sqrt{3}\sigma_z\big)}{2}\,, \quad (2.16.1)$$

which seems to tell us that ρ is obtained by putting together a fraction of $(3+\sqrt{3})/6 = 78.9\%$ of atoms that are up in the direction of the unit vector with coordinates $\frac{1}{2}\begin{pmatrix} 1 \\ 0 \\ \sqrt{3} \end{pmatrix}$ and a fraction of $(3-\sqrt{3})/6 = 21.1\%$ of atoms that are up in the direction of the unit vector with coordinates $\frac{1}{2}\begin{pmatrix} 1 \\ 0 \\ -\sqrt{3} \end{pmatrix}$. It is the same statistical operator ρ in both cases, although the preparation procedures are quite different. And, clearly, if you have two different ways of blending one and the same mixture, you have many.

The last sentence introduces some terminology. In expressions such as

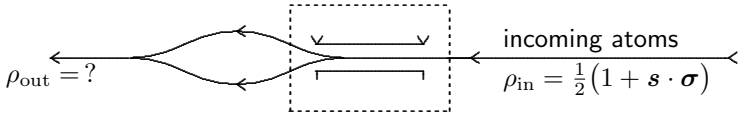
$$\rho = \frac{1}{2}\left(1 + \frac{1}{2}\sigma_x + \frac{1}{2}\sigma_z\right) = \frac{1}{2}\frac{1+\sigma_x}{2} + \frac{1}{2}\frac{1+\sigma_z}{2}\,, \quad (2.16.2)$$

we call the weighted sum of projectors on the right a *blend* of the *mixture* on the left. As we have seen, there are many ways, as a rule, to blend any given mixture. Can you tell, by looking at the mixture, how it was blended from ingredients? No, you cannot, because all statistical data you can get are expectation values in accordance with the Born rule of (2.15.2), $\langle\Lambda\rangle = \mathrm{tr}(\Lambda\rho)$, and for them only the mixture ρ is relevant, not how it is blended. In the example of (2.16.2), it is all right to say that the mixture is "as if 50% are $\uparrow_x$ and 50% are $\uparrow_z$," but this is only one *as-if reality* of many.

2.17 Nonselective measurement

Having statistical operators at our disposal for the characterization of approaching atoms, we can finally close a gap, namely how do we describe, in the mathematical formalism, the procedure of measuring σ_z, say, without

however selecting a partial beam? The situation is as follows:



(2.17.1)

Clearly, the emerging atoms are here blended from $\uparrow_z$ and $\downarrow_z$, and the weights are the probabilities for an incoming atom to be deflected correspondingly. Accordingly, we have

$$\rho_{\rm out} = \underbrace{\frac{1}{2}(1+s_z)}_{\substack{\text{probability}\\ \text{for}\uparrow_z}}\underbrace{\frac{1}{2}(1+\sigma_z)}_{\substack{\text{projector}\\ \text{on}\uparrow_z}} + \underbrace{\frac{1}{2}(1-s_z)}_{\substack{\text{probability}\\ \text{for}\downarrow_z}}\underbrace{\frac{1}{2}(1-\sigma_z)}_{\substack{\text{projector}\\ \text{on}\downarrow_z}} \tag{2.17.2}$$

or

$$\rho_{\rm out} = \frac{1}{2}(1 + s_z\sigma_z)\,. \tag{2.17.3}$$

The comparison with

$$\rho_{\rm in} = \frac{1}{2}(1 + s_x\sigma_x + s_y\sigma_y + s_z\sigma_z) \tag{2.17.4}$$

shows that the nonselective measurement effectively wipes out the components referring to the perpendicular directions:

$$\boldsymbol{s} \mathrel{\widehat{=}} \underbrace{\begin{pmatrix} s_x \\ s_y \\ s_z \end{pmatrix}}_{\text{in}} \longrightarrow \boldsymbol{s} \mathrel{\widehat{=}} \underbrace{\begin{pmatrix} 0 \\ 0 \\ s_z \end{pmatrix}}_{\text{out}}\,. \tag{2.17.5}$$

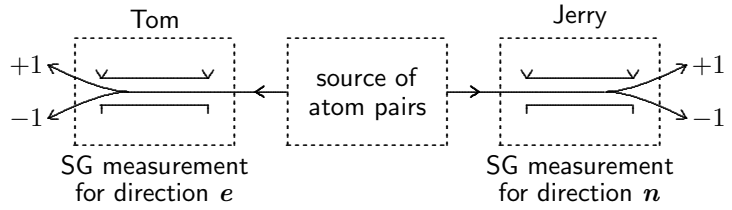
See also Exercise 44.

It does not matter, of course, whether we (or someone else) cares to keep a record of which atom was deflected which way because this information is not taken into account anyway. What enters into the statistical operator is solely the information upon which we condition our predictions. In the above example, this is as follows: Provided that the source emits atoms in state $\rho_{\rm in}$ and that they traverse a z-measuring Stern–Gerlach apparatus without being selected, what is then the expectation value of Λ for the beam thus prepared? Answer:

$$\langle\Lambda\rangle = {\rm tr}(\Lambda\rho_{\rm out}) = \frac{1}{2}{\rm tr}(\Lambda) + \frac{1}{2}s_z{\rm tr}(\Lambda\sigma_z)\,. \tag{2.17.6}$$

2.18 Entangled atom pairs

So far we have been dealing exclusively with the quantum aspects of the magnetic properties of single silver atoms. Let us now move on and turn to more complicated systems. For a start, we consider pairs of silver atoms in a situation much like that discussed in Section 1.2 for paired photons. We depict the analog of (1.2.1):



(2.18.1)

The source emits pairs of atoms. One travels to the left and the other to the right. On the left, where *Tom* is taking data, his atom passes through a Stern–Gerlach magnet, set for probing direction $\boldsymbol{e}$, and number $+1$ or -1 is recorded for each atom. Likewise, on the right, *Jerry* orients his apparatus to probe direction $\boldsymbol{n}$ and records $+1$ or -1 for each atom.

For a particular choice of direction, they eventually calculate the Bell correlation $C(\boldsymbol{e},\boldsymbol{n})$ of (1.2.3) in accordance with

$$\begin{aligned} C(\boldsymbol{e},\boldsymbol{n}) = {} & (\text{relative frequency of products} = +1) \\ & - (\text{relative frequency of products} = -1)\,, \end{aligned} \tag{2.18.2}$$

or expressed in terms of probabilities for $\uparrow\uparrow$, $\uparrow\downarrow$, $\downarrow\uparrow$, $\downarrow\downarrow$,

$$C(\boldsymbol{e},\boldsymbol{n}) = \text{prob}(\uparrow\uparrow \text{ or } \downarrow\downarrow) - \text{prob}(\uparrow\downarrow \text{ or } \downarrow\uparrow). \tag{2.18.3}$$

Clearly, the value of $C(\boldsymbol{e},\boldsymbol{n})$ depends on the directions that are probed and also on the two-atom state emitted by the source.

The choice of directions enters the Bell correlation $C(\boldsymbol{e},\boldsymbol{n})$ in the form of projection operators $\frac{1}{2}\big(1 \pm \boldsymbol{e}\cdot\boldsymbol{\sigma}^{(1)}\big)$ and $\frac{1}{2}\big(1 \pm \boldsymbol{n}\cdot\boldsymbol{\sigma}^{(2)}\big)$, where $\boldsymbol{\sigma}^{(1)}$ is the Pauli vector operator for Tom's atom on the left and $\boldsymbol{\sigma}^{(2)}$ is that for Jerry's atom on the right. The respective probabilities are then the expectation

values of the corresponding products of projection operators,

$$\begin{aligned}
\text{prob}(\uparrow\uparrow) &= \left\langle \frac{1+\boldsymbol{e}\cdot\boldsymbol{\sigma}^{(1)}}{2}\frac{1+\boldsymbol{n}\cdot\boldsymbol{\sigma}^{(2)}}{2}\right\rangle,\\
\text{prob}(\downarrow\downarrow) &= \left\langle \frac{1-\boldsymbol{e}\cdot\boldsymbol{\sigma}^{(1)}}{2}\frac{1-\boldsymbol{n}\cdot\boldsymbol{\sigma}^{(2)}}{2}\right\rangle,\\
\text{prob}(\uparrow\downarrow) &= \left\langle \frac{1+\boldsymbol{e}\cdot\boldsymbol{\sigma}^{(1)}}{2}\frac{1-\boldsymbol{n}\cdot\boldsymbol{\sigma}^{(2)}}{2}\right\rangle,\\
\text{prob}(\downarrow\uparrow) &= \left\langle \frac{1-\boldsymbol{e}\cdot\boldsymbol{\sigma}^{(1)}}{2}\frac{1+\boldsymbol{n}\cdot\boldsymbol{\sigma}^{(2)}}{2}\right\rangle,
\end{aligned} \tag{2.18.4}$$

where, for instance, prob(↑↓) is the probability that Tom's atom is found as ↑ and Jerry's as ↓. Accordingly,

$$\begin{aligned}
\text{prob}(\uparrow\uparrow \text{ or } \downarrow\downarrow) &= \text{prob}(\uparrow\uparrow) + \text{prob}(\downarrow\downarrow)\\
&= \left\langle \frac{1}{2}\Big(1+\boldsymbol{e}\cdot\boldsymbol{\sigma}^{(1)}\ \boldsymbol{n}\cdot\boldsymbol{\sigma}^{(2)}\Big)\right\rangle
\end{aligned} \tag{2.18.5}$$

and

$$\begin{aligned}
\text{prob}(\uparrow\downarrow \text{ or } \downarrow\uparrow) &= \text{prob}(\uparrow\downarrow) + \text{prob}(\downarrow\uparrow)\\
&= \left\langle \frac{1}{2}\Big(1-\boldsymbol{e}\cdot\boldsymbol{\sigma}^{(1)}\ \boldsymbol{n}\cdot\boldsymbol{\sigma}^{(2)}\Big)\right\rangle
\end{aligned} \tag{2.18.6}$$

so that

$$C(\boldsymbol{e},\boldsymbol{n}) = \left\langle \boldsymbol{e}\cdot\boldsymbol{\sigma}^{(1)}\ \boldsymbol{n}\cdot\boldsymbol{\sigma}^{(2)}\right\rangle = \text{tr}\Big(\boldsymbol{e}\cdot\boldsymbol{\sigma}^{(1)}\ \boldsymbol{n}\cdot\boldsymbol{\sigma}^{(2)}\rho\Big), \tag{2.18.7}$$

where ρ is the statistical operator that specifies the state of the atom pairs emitted by the source.

The source could, for example, emit the atoms in a state with the ket $|\uparrow_z\downarrow_z\rangle$, that is, Tom's atom is $\uparrow_z$ and Jerry's atom is $\downarrow_z$. Or in a state with the ket $|\downarrow_z\uparrow_z\rangle$, now Tom's atom is $\downarrow_z$ and Jerry's atom is $\uparrow_z$. Of course, there are also the possibilities of $|\uparrow_z\uparrow_z\rangle$, or $|\downarrow_z\downarrow_z\rangle$, or any other ket comprised of them as a weighted sum,

$$\big|\ \big\rangle = \big|\uparrow_z\uparrow_z\big\rangle\alpha + \big|\downarrow_z\downarrow_z\big\rangle\beta + \big|\uparrow_z\downarrow_z\big\rangle\gamma + \big|\downarrow_z\uparrow_z\big\rangle\delta\,, \tag{2.18.8}$$

where the probability amplitudes $\alpha,\ldots,\delta$ are normalized in accordance with

$$|\alpha|^2 + |\beta|^2 + |\gamma|^2 + |\delta|^2 = 1\,. \tag{2.18.9}$$

These are simple self-suggesting extensions of the formalism that we have developed for single atoms. The main difference is that we now have four basic options, two for each atom, and we need to convey more information when labeling the kets. Correspondingly, an orthonormal basis for the kets now requires four kets such as those on the right-hand side in (2.18.8), and a basis for the bras has four bras.

The basis ket $|{\uparrow_z}{\downarrow_z}\rangle$ refers to the situation in which Tom's atom by itself is described by $|{\uparrow_z}\rangle$ and Jerry's atom by itself by $|{\downarrow_z}\rangle$. Such a ket is a *tensor product* of its ingredients,

$$\big|{\uparrow_z}{\downarrow_z}\big\rangle = \big|{\uparrow_z}\big\rangle \otimes \big|{\downarrow_z}\big\rangle\,, \tag{2.18.10}$$

with the two-party ket regarded as a product of two single-party kets. The adjoint statement about the bras,

$$\big\langle{\uparrow_z}{\downarrow_z}\big| = \big\langle{\uparrow_z}\big| \otimes \big\langle{\downarrow_z}\big|\,, \tag{2.18.11}$$

is natural and so are brackets,

$$\langle a,b|a',b'\rangle = \Big(\langle a|\otimes\langle b|\Big)\Big(|a'\rangle\otimes|b'\rangle\Big) = \langle a|a'\rangle\langle b|b'\rangle\,, \tag{2.18.12}$$

the bracket of two tensor products is the product of the single-party brackets.

Analogously, we have tensor products for ket-bras,

$$\Big(|a'\rangle\langle a|\Big)\otimes\Big(|b'\rangle\langle b|\Big) = |a',b'\rangle\langle a,b| = \Big(|a'\rangle\otimes|b'\rangle\Big)\otimes\Big(\langle a|\otimes\langle b|\Big)\,. \tag{2.18.13}$$

The Pauli operators $\boldsymbol{\sigma}^{(1)}$ and $\boldsymbol{\sigma}^{(2)}$ for Tom's and Jerry's atoms, respectively, are tensor products with the identity for the other atom,

$$\boldsymbol{\sigma}^{(1)} = \boldsymbol{\sigma}\otimes 1 \quad\text{and}\quad \boldsymbol{\sigma}^{(2)} = 1\otimes\boldsymbol{\sigma}\,, \tag{2.18.14}$$

if we emphasize the tensor-product structure. The product in (2.18.7), then, is

$$\boldsymbol{e}\cdot\boldsymbol{\sigma}^{(1)}\;\boldsymbol{n}\cdot\boldsymbol{\sigma}^{(2)} = \boldsymbol{e}\cdot\boldsymbol{\sigma}\otimes\boldsymbol{n}\cdot\boldsymbol{\sigma}\,, \tag{2.18.15}$$

where we prefer the left-hand side as the tensor-product notation is a bit tiring.

More generally than (2.18.10), we have

$$\begin{aligned}&\underbrace{\Big(\big|{\uparrow_z}\big\rangle\alpha_1 + \big|{\downarrow_z}\big\rangle\beta_1\Big)}_{\text{Tom's atom}}\otimes\underbrace{\Big(\big|{\uparrow_z}\big\rangle\alpha_2 + \big|{\downarrow_z}\big\rangle\beta_2\Big)}_{\text{Jerry's atom}}\\ &= \big|{\uparrow_z}{\uparrow_z}\big\rangle\alpha_1\alpha_2 + \big|{\downarrow_z}{\downarrow_z}\big\rangle\beta_1\beta_2 + \big|{\uparrow_z}{\downarrow_z}\big\rangle\alpha_1\beta_2 + \big|{\downarrow_z}{\uparrow_z}\big\rangle\beta_1\alpha_2\,.\end{aligned} \tag{2.18.16}$$

This two-party product ket is a particular case of (2.18.8) with

$$\alpha\beta = \gamma\delta\,. \tag{2.18.17}$$

This restriction on the amplitudes in (2.18.8) is both necessary and sufficient for the two-party ket to be a tensor product of two single-party kets; see Exercise 45.

The generic two-party ket in (2.18.8), however, is not a product ket, it is a superposition of product kets for which (2.18.17) does not hold. If this is the case, we do not have kets for Tom's and Jerry's atoms by themselves, we have a ket for the two-atom system. In this situation, the individual atoms have no properties of their own, while the pair of atoms has distinct properties. Such an intimate link is called *entanglement*, a term coined by Schrödinger. One says the magnetic degrees of freedom of the two atoms are entangled or, more sloppily, that the atoms are entangled. Terminology aside, what is important here is that an atom pair whose properties are described by a ket (2.18.8) with $\alpha\beta \neq \gamma\delta$, a ket of the entangled kind, is one single physical object, not two objects — one entangled pair of atoms, not two individual atoms.

Picking up the story at (2.18.7), we are aiming at establishing a situation in which the Bell inequality (1.2.10) is violated by quantum correlations that are substantially stronger than any classical correlation could ever be. For this purpose, it is simplest to consider a source that emits the two-atom state described by the ket

$$|\,\rangle = \Big(|{\uparrow_z}{\downarrow_z}\rangle - |{\downarrow_z}{\uparrow_z}\rangle\Big)\Big/\sqrt{2}\,, \tag{2.18.18}$$

clearly an entangled state as $0 \neq -\frac{1}{2}$ in (2.18.17). As noted above, the atoms have no individual properties in a state of this kind. Jointly, the two atoms have the distinct objective property that if one is $\uparrow_z$, the other will be $\downarrow_z$. Which one is $\uparrow_z$, and which is $\downarrow_z$, is not predictable but that they are paired in this fashion is. Put differently, if Tom chooses $\boldsymbol{e} = \boldsymbol{e}_z$ and Jerry chooses $\boldsymbol{n} = \boldsymbol{e}_z$, the measurement results will either be $+1$ and -1 or -1 and $+1$ but never $+1$ and $+1$ or -1 and -1. For this setting, then, we have

$$C(\boldsymbol{e}, \boldsymbol{n}) = -1 \qquad \text{if} \quad \boldsymbol{e} = \boldsymbol{n} = \boldsymbol{e_z} \mathrel{\widehat{=}} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}. \tag{2.18.19}$$

The statistical operator $\rho = |\,\rangle\langle\,|$ for this two-atom state is

$$\begin{aligned}\rho^{(\mathrm{pair})} &= \Big(\big|{\uparrow_z}{\downarrow_z}\big\rangle - \big|{\downarrow_z}{\uparrow_z}\big\rangle\Big)\frac{1}{2}\Big(\big\langle{\uparrow_z}{\downarrow_z}\big| - \big\langle{\downarrow_z}{\uparrow_z}\big|\Big) \\ &= \big|{\uparrow_z}{\downarrow_z}\big\rangle\frac{1}{2}\big\langle{\uparrow_z}{\downarrow_z}\big| + \big|{\downarrow_z}{\uparrow_z}\big\rangle\frac{1}{2}\big\langle{\downarrow_z}{\uparrow_z}\big| \\ &\quad - \big|{\uparrow_z}{\downarrow_z}\big\rangle\frac{1}{2}\big\langle{\downarrow_z}{\uparrow_z}\big| - \big|{\downarrow_z}{\uparrow_z}\big\rangle\frac{1}{2}\big\langle{\uparrow_z}{\downarrow_z}\big|\,, \end{aligned} \tag{2.18.20}$$

which we express in terms of $\boldsymbol{\sigma}^{(1)}$ and $\boldsymbol{\sigma}^{(2)}$ by first noting that, for each atom by itself,

$$\left.\begin{matrix}\big|{\uparrow_z}\big\rangle\big\langle{\uparrow_z}\big| \\ \big|{\downarrow_z}\big\rangle\big\langle{\downarrow_z}\big|\end{matrix}\right\} = \frac{1}{2}(1 \pm \sigma_z)\,,$$
$$\left.\begin{matrix}\big|{\uparrow_z}\big\rangle\big\langle{\downarrow_z}\big| \\ \big|{\downarrow_z}\big\rangle\big\langle{\uparrow_z}\big|\end{matrix}\right\} = \frac{1}{2}(\sigma_x \pm \mathrm{i}\sigma_y)\,. \tag{2.18.21}$$

Thus,

$$\big|{\uparrow_z}{\downarrow_z}\big\rangle\big\langle{\uparrow_z}{\downarrow_z}\big| = \frac{1}{2}\big(1 + \sigma_z^{(1)}\big)\frac{1}{2}\big(1 - \sigma_z^{(2)}\big)\,, \tag{2.18.22}$$

for instance, and we get

$$\begin{aligned}\rho^{(\mathrm{pair})} &= \frac{1}{2}\frac{1+\sigma_z^{(1)}}{2}\frac{1-\sigma_z^{(2)}}{2} + \frac{1}{2}\frac{1-\sigma_z^{(1)}}{2}\frac{1+\sigma_z^{(2)}}{2} \\ &\quad - \frac{1}{2}\frac{\sigma_x^{(1)} + \mathrm{i}\sigma_y^{(1)}}{2}\frac{\sigma_x^{(2)} - \mathrm{i}\sigma_y^{(2)}}{2} - \frac{1}{2}\frac{\sigma_x^{(1)} - \mathrm{i}\sigma_y^{(1)}}{2}\frac{\sigma_x^{(2)} + \mathrm{i}\sigma_y^{(2)}}{2} \\ &= \frac{1}{4}\big(1 - \sigma_x^{(1)}\sigma_x^{(2)} - \sigma_y^{(1)}\sigma_y^{(2)} - \sigma_z^{(1)}\sigma_z^{(2)}\big) \\ &= \frac{1}{4}\big(1 - \boldsymbol{\sigma}^{(1)}\cdot\boldsymbol{\sigma}^{(2)}\big)\,. \end{aligned} \tag{2.18.23}$$

In the tensor-product notation of (2.18.15), this appears as

$$\rho^{(\mathrm{pair})} = \frac{1}{4}\big(1 - \sigma_x\otimes\sigma_x - \sigma_y\otimes\sigma_y - \sigma_z\otimes\sigma_z\big)\,, \tag{2.18.24}$$

where we do not have the notational convenience of the scalar product $\boldsymbol{\sigma}^{(1)}\cdot\boldsymbol{\sigma}^{(2)}$ unless we invent a multiplication sign that combines $\cdot$ and $\otimes$.

Note two things about this statistical operator. First, the identity is here multiplied by $\frac{1}{4}$ because we now have a four-dimensional ket space (and bra space) so that the matrix representation is by 4×4 matrices and $\mathrm{tr}(1) = 4$; in tensor notation, the two-party identity is the product of two single-party identities: $1 = 1\otimes 1$.

Second, there is no difference between x, y, and z in the compact final form of $\rho^{(\text{pair})}$ in (2.18.23), which implies immediately that we also have

$$C(\boldsymbol{e},\boldsymbol{n}) = -1 \quad \text{if} \quad \boldsymbol{e} = \boldsymbol{n} = \boldsymbol{e}_x \mathrel{\widehat{=}} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \quad \text{or} \quad \boldsymbol{e} = \boldsymbol{n} = \boldsymbol{e}_y \mathrel{\widehat{=}} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$
$$\text{or} \quad \boldsymbol{e} = \boldsymbol{n} \text{ in any direction.} \tag{2.18.25}$$

Despite the original reference to the z direction in (2.18.18), no spatial direction is singled out by this two-atom state.

Now, the trace in

$$C(\boldsymbol{e},\boldsymbol{n}) = \text{tr}\Big(\boldsymbol{e}\cdot\boldsymbol{\sigma}^{(1)}\ \boldsymbol{n}\cdot\boldsymbol{\sigma}^{(2)}\frac{1}{4}\big(1-\boldsymbol{\sigma}^{(1)}\cdot\boldsymbol{\sigma}^{(2)}\big)\Big) \tag{2.18.26}$$

refers to both atoms,

$$\text{tr}(\ ) = \text{tr}_{\text{pair}}(\ ) = \text{tr}_1\big(\text{tr}_2(\ )\big) = \text{tr}_2\big(\text{tr}_1(\ )\big)\,, \tag{2.18.27}$$

and we evaluate it successively

$$C(\boldsymbol{e},\boldsymbol{n}) = \text{tr}_1\Big(\boldsymbol{e}\cdot\boldsymbol{\sigma}^{(1)}\ \text{tr}_2\Big(\boldsymbol{n}\cdot\boldsymbol{\sigma}^{(2)}\frac{1}{4}\big(1-\boldsymbol{\sigma}^{(1)}\cdot\boldsymbol{\sigma}^{(2)}\big)\Big)\Big). \tag{2.18.28}$$

Upon recalling from (2.15.10) and (2.15.20) that, for a single atom,

$$\text{tr}(\boldsymbol{a}\cdot\boldsymbol{\sigma}) = 0\,, \qquad \text{tr}(\boldsymbol{a}\cdot\boldsymbol{\sigma}\ \ \boldsymbol{b}\cdot\boldsymbol{\sigma}) = 2\boldsymbol{a}\cdot\boldsymbol{b}\,, \tag{2.18.29}$$

and noting that $\boldsymbol{\sigma}^{(1)}$ is just like a numerical vector for atom 2, we have

$$\text{tr}_2\Big(\boldsymbol{n}\cdot\boldsymbol{\sigma}^{(2)}\frac{1}{4}\big(1-\boldsymbol{\sigma}^{(1)}\cdot\boldsymbol{\sigma}^{(2)}\big)\Big) = -\frac{1}{2}\boldsymbol{n}\cdot\boldsymbol{\sigma}^{(1)} \tag{2.18.30}$$

and then

$$C(\boldsymbol{e},\boldsymbol{n}) = \text{tr}_1\Big(\boldsymbol{e}\cdot\boldsymbol{\sigma}^{(1)}\Big(-\frac{1}{2}\boldsymbol{n}\cdot\boldsymbol{\sigma}^{(1)}\Big)\Big) = -\boldsymbol{e}\cdot\boldsymbol{n}\,, \tag{2.18.31}$$

fully consistent with the observation in (2.18.25) that $C(\boldsymbol{e},\boldsymbol{n}) = -1$ for $\boldsymbol{e} = \boldsymbol{n} =$ any direction.

The left-hand side of the Bell inequality (1.2.10) — now with $a \to \boldsymbol{e}_1$, $a' \to \boldsymbol{e}_2$ and $b \to \boldsymbol{n}_1$, $b' \to \boldsymbol{n}_2$ as fits the present situation in which parameter settings are choices of directions — is then

$$\begin{aligned} L &= \big|C(\boldsymbol{e}_1,\boldsymbol{n}_1) + C(\boldsymbol{e}_1,\boldsymbol{n}_2) + C(\boldsymbol{e}_2,\boldsymbol{n}_1) - C(\boldsymbol{e}_2,\boldsymbol{n}_2)\big| \\ &= \big|\boldsymbol{e}_1\cdot(\boldsymbol{n}_1+\boldsymbol{n}_2) + \boldsymbol{e}_2\cdot(\boldsymbol{n}_1-\boldsymbol{n}_2)\big|\,, \end{aligned} \tag{2.18.32}$$

and we wish to choose the four directions such that we maximize the value of L. We do this in two steps. First, for given $\boldsymbol{n}_1$ and $\boldsymbol{n}_2$, we optimize $\boldsymbol{e}_1$ and $\boldsymbol{e}_2$ clearly by having $\boldsymbol{e}_1$ in the direction of the sum vector $\boldsymbol{n}_1 + \boldsymbol{n}_2$ and $\boldsymbol{e}_2$ in the direction of the difference vector $\boldsymbol{n}_1 - \boldsymbol{n}_2$,

$$\boldsymbol{e}_1 = \frac{\boldsymbol{n}_1 + \boldsymbol{n}_2}{|\boldsymbol{n}_1 + \boldsymbol{n}_2|} \qquad \text{and} \qquad \boldsymbol{e}_2 = \frac{\boldsymbol{n}_1 - \boldsymbol{n}_2}{|\boldsymbol{n}_1 - \boldsymbol{n}_2|}\,, \tag{2.18.33}$$

which are orthogonal, $\boldsymbol{e}_1 \cdot \boldsymbol{e}_2 = 0$, so that

$$\begin{aligned} \operatorname*{Max}_{\boldsymbol{e}_{1,2}}\{L\} &= \big|\boldsymbol{n}_1 + \boldsymbol{n}_2\big| + \big|\boldsymbol{n}_1 - \boldsymbol{n}_2\big| \\ &= \sqrt{2 + 2\boldsymbol{n}_1 \cdot \boldsymbol{n}_2} + \sqrt{2 - 2\boldsymbol{n}_1 \cdot \boldsymbol{n}_2}\,. \end{aligned} \tag{2.18.34}$$

Upon denoting the angle between $\boldsymbol{n}_1$ and $\boldsymbol{n}_2$ by θ, $\boldsymbol{n}_1 \cdot \boldsymbol{n}_2 = \cos(\theta)$ with $0 \le \theta \le \pi$, this reads

$$\begin{aligned} \operatorname*{Max}_{\boldsymbol{e}_{1,2}}\{L\} &= 2\sin\left(\frac{\theta}{2}\right) + 2\cos\left(\frac{\theta}{2}\right) \\ &= 2\sqrt{2}\left(\underbrace{\frac{1}{\sqrt{2}}}_{=\cos(\pi/4)}\cos\left(\frac{\theta}{2}\right) + \underbrace{\frac{1}{\sqrt{2}}}_{=\sin(\pi/4)}\sin\left(\frac{\theta}{2}\right)\right) \\ &= 2\sqrt{2}\,\cos\left(\frac{\theta}{2} - \frac{\pi}{4}\right). \end{aligned} \tag{2.18.35}$$

The largest value is obtained for $\theta = \frac{1}{2}\pi$ when $\boldsymbol{n}_1 \cdot \boldsymbol{n}_2 = 0$ in (2.18.34) and

$$\operatorname{Max}\{L\} = 2\sqrt{2} \tag{2.18.36}$$

clearly exceeds the upper bound of the Bell inequality (1.2.10), $L \le 2$. We conclude that the Bell inequality does not apply to quantum correlations of the type built into entangled two-atom states of the kind exemplified by (2.18.18) where the individual atoms have no properties of their own, while the pair of atoms has distinct statistical properties.

This closes the argument given in Section 1.2 and ultimately justifies the statement (1.2.11) about the nonexistence of any mechanism that would decide the outcome of a quantum measurement. Owing to the experiments pioneered by Clauser* and Aspect,† and perfected by Zeilinger‡ and others, the violation of the Bell inequality is an established fact. Nowadays, it is a routine matter to check that $2 < L \le 2\sqrt{2}$ as part of the calibration

*John Francis CLAUSER (b. 1942) †Alain ASPECT (b. 1947)
‡Anton ZEILINGER (b. 1945)

procedure in experiments with entangled degrees of freedom. In summary, the Bell inequality is wrong as a statement about correlations observed in quantum experiments.

2.19 State reduction, conditional probabilities

Another point worth mentioning is the following. Suppose Tom is unaware of Jerry and his measurements and just sees the atoms emitted by the source to the left. They are a mixture of $\uparrow_z$ and $\downarrow_z$, or of $\uparrow_x$ and $\downarrow_x$, or ..., to him,

$$\begin{aligned}\rho_{\mathrm{T}}^{(1)} &= \frac{1}{2}\frac{1+\sigma_z^{(1)}}{2} + \frac{1}{2}\frac{1-\sigma_z^{(1)}}{2} = \frac{1}{2} \\ &= \frac{1}{2}\frac{1+\sigma_x^{(1)}}{2} + \frac{1}{2}\frac{1-\sigma_x^{(1)}}{2} = \frac{1}{2}\,,\end{aligned} \tag{2.19.1}$$

and Tom makes statistical predictions about the next atom consistent with this statistical operator. How about Jerry, then? That depends. As long as he has not measured his atom on the right, Jerry also uses

$$\rho_{\mathrm{J}}^{(1)} = \frac{1}{2} \tag{2.19.2}$$

to predict the statistics of measurement results for Tom's atom. But after he measured his atom along z and found $\uparrow_z$, say, then Jerry uses

$$\rho_{\mathrm{J}}^{(1)} = \frac{1-\sigma_z^{(1)}}{2} \tag{2.19.3}$$

since he knows that Tom's atom is $\downarrow_z$ under these circumstances. And if he measured his atom along x and found $\downarrow_x$, say, then Jerry uses

$$\rho_{\mathrm{J}}^{(1)} = \frac{1+\sigma_x^{(1)}}{2} \tag{2.19.4}$$

because he knows that Tom's atom has to be $\uparrow_x$ under these circumstances. The transitions

$$\rho_{\mathrm{J}}^{(1)} = \frac{1}{2} \to \begin{cases} \rho_{\mathrm{J}}^{(1)} = \dfrac{1-\sigma_z^{(1)}}{2} \\ \rho_{\mathrm{J}}^{(1)} = \dfrac{1+\sigma_x^{(1)}}{2} \end{cases} \tag{2.19.5}$$

are examples of *state reduction*, a bookkeeping tool — part and parcel of all statistical formalisms — by which additional, newly acquired knowledge is taken into account in order to update our statistical predictions accordingly.

State reduction is not a physical process; nothing happens to Tom's atom when Jerry measures his. State reduction is a mental process reflecting that Jerry has identified Tom's atom as a member of a particular subensemble.

It is thus possible that Tom uses

$$\rho_{\mathrm{T}}^{(1)} = \frac{1}{2} \tag{2.19.6}$$

and Jerry uses

$$\rho_{\mathrm{J}}^{(1)} = \frac{1 - \sigma_z^{(1)}}{2} \tag{2.19.7}$$

for statistical predictions about one and the same next atom, and both make statistically correct predictions. This teaches us the lesson that the statistical operator is not a property of the atom, such as its mass or its electric charge, but a theoretical tool of the physicist who talks about the atom professionally, thereby taking into account all *he* knows about the atom. Another physicist may have different knowledge, and then she will use another statistical operator, one that correctly reflects *her* knowledge.

We recognize here that the probabilities that we calculate with the quantum formalism are always conditional probabilities, they are conditioned on what we know about the situation. The same conditioning enters when we compute expectation values in accordance with the Born rule in (2.15.2), $\langle \Lambda \rangle = \mathrm{tr}(\Lambda\rho)$. The operator Λ represents the property in question and all parties involved (such as Tom and Jerry) agree on its form. In marked contrast, our statistical operator represents what *we* know about the preparation of the physical system, and our knowledge can very well be different from yours, in which case you will use a statistical operator different from ours. Although you and we may get different values for $\langle \Lambda \rangle$, both values are correct; they refer to different circumstances.

Since the statistical operator we use encodes what we know about the circumstances, we have to update it when we acquire new information and the circumstances change. Typically, the new information results from the outcome of a measurement. State reduction is the tool that converts the previous, out-dated statistical operator into the new one.

In more general situations, we may not even have enough information to specify a statistical operator as the proper representation of our knowledge about the quantum system. Then, there is a list of possible ρs with probabilities of occurrence. The assignment of these probabilities in view of the observed data requires the tools of quantum state estimation, which is a topic beyond the ground covered here.

2.20 Measurement outcomes do not pre-exist

In Section 2.18, we considered expectation values of operators composed of tensor products of one component each of the Pauli vector operators for Tom's and Jerry's atoms, such as

$$\boldsymbol{e}\cdot\boldsymbol{\sigma}^{(1)}\,\boldsymbol{n}\cdot\boldsymbol{\sigma}^{(2)} = \boldsymbol{e}\cdot\boldsymbol{\sigma}\otimes\boldsymbol{n}\cdot\boldsymbol{\sigma}\,. \tag{2.20.1}$$

Together, there are nine basic atom-pair operators of this kind. Following Mermin's* suggestion, we arrange them in a 3×3 table,

$$\begin{array}{ccc} \sigma_x^{(1)}\sigma_x^{(2)} & \sigma_y^{(1)}\sigma_y^{(2)} & \sigma_z^{(1)}\sigma_z^{(2)} \\ \sigma_y^{(1)}\sigma_z^{(2)} & \sigma_z^{(1)}\sigma_x^{(2)} & \sigma_x^{(1)}\sigma_y^{(2)} \\ \sigma_z^{(1)}\sigma_y^{(2)} & \sigma_x^{(1)}\sigma_z^{(2)} & \sigma_y^{(1)}\sigma_x^{(2)} \end{array} \tag{2.20.2}$$

where the three operators in each row and each column commute with each other. For example, in the top row, we have

$$\begin{aligned} &\left(\sigma_x^{(1)}\sigma_x^{(2)}\right)\left(\sigma_y^{(1)}\sigma_y^{(2)}\right) = \left(\sigma_x^{(1)}\sigma_y^{(1)}\right)\left(\sigma_x^{(2)}\sigma_y^{(2)}\right) \\ &= \left(-\sigma_y^{(1)}\sigma_x^{(1)}\right)\left(-\sigma_y^{(2)}\sigma_x^{(2)}\right) = \left(\sigma_y^{(1)}\sigma_y^{(2)}\right)\left(\sigma_x^{(1)}\sigma_x^{(2)}\right), \end{aligned} \tag{2.20.3}$$

where the occurrence of two minus signs, with no net effect, is the crucial observation. This happens for all products of two table entries from the same row or the same column.

We note also that the product of any two same-row or same-column entries in Mermin's table (2.20.2) equals the third entry or its negative. As an example, consider the middle column,

$$\begin{aligned} \left(\sigma_y^{(1)}\sigma_y^{(2)}\right)\left(\sigma_z^{(1)}\sigma_x^{(2)}\right) &= \left(\sigma_y^{(1)}\sigma_z^{(1)}\right)\left(\sigma_y^{(2)}\sigma_x^{(2)}\right) \\ &= \left(\mathrm{i}\sigma_x^{(1)}\right)\left(-\mathrm{i}\sigma_z^{(2)}\right) = \sigma_x^{(1)}\sigma_z^{(2)}\,, \end{aligned} \tag{2.20.4}$$

or the middle row,

$$\begin{aligned} \left(\sigma_y^{(1)}\sigma_z^{(2)}\right)\left(\sigma_z^{(1)}\sigma_x^{(2)}\right) &= \left(\sigma_y^{(1)}\sigma_z^{(1)}\right)\left(\sigma_z^{(2)}\sigma_x^{(2)}\right) \\ &= \left(\mathrm{i}\sigma_x^{(1)}\right)\left(\mathrm{i}\sigma_y^{(2)}\right) = -\sigma_x^{(1)}\sigma_y^{(2)}\,. \end{aligned} \tag{2.20.5}$$

Since the square of each table entry is the two-atom identity operator, we get ± 1 for the products of all entries in each row or each column; in fact,

*Nathaniel David MERMIN (b. 1935)

it is $+1$ for the column products and -1 for the row products (Exercise 21 helps here). We supplement (2.20.2) with these products,

$$\begin{array}{ccc|c} \sigma_x^{(1)}\sigma_x^{(2)} & \sigma_y^{(1)}\sigma_y^{(2)} & \sigma_z^{(1)}\sigma_z^{(2)} & -1 \\ \sigma_y^{(1)}\sigma_z^{(2)} & \sigma_z^{(1)}\sigma_x^{(2)} & \sigma_x^{(1)}\sigma_y^{(2)} & -1 \\ \sigma_z^{(1)}\sigma_y^{(2)} & \sigma_x^{(1)}\sigma_z^{(2)} & \sigma_y^{(1)}\sigma_x^{(2)} & -1 \\ \hline +1 & +1 & +1 & \end{array} \tag{2.20.6}$$

and are now ready for the question we want to ask and answer.

Question: Do physical properties have the observed values before the measurement?
Answer: No, they don't. (2.20.7)

For Mermin's table, the question is about the values of the nine atom-pair observables, each with two eigenvalues $+1$ and two eigenvalues -1. Taking the top row as an example once more, the four kets

$$\begin{aligned} |---\rangle &= \Big(|\uparrow_z\downarrow_z\rangle - |\downarrow_z\uparrow_z\rangle\Big)\Big/\sqrt{2}\,, \\ |-++\rangle &= \Big(|\uparrow_z\uparrow_z\rangle - |\downarrow_z\downarrow_z\rangle\Big)\Big/\sqrt{2}\,, \\ |+-+\rangle &= \Big(|\uparrow_z\uparrow_z\rangle + |\downarrow_z\downarrow_z\rangle\Big)\Big/\sqrt{2}\,, \\ |++-\rangle &= \Big(|\uparrow_z\downarrow_z\rangle + |\downarrow_z\uparrow_z\rangle\Big)\Big/\sqrt{2} \end{aligned} \tag{2.20.8}$$

are an orthonormal basis and each is an eigenstate of the top-row operators, where the notation indicates the respective eigenvalues. Consistent with (2.20.6), the product of the three eigenvalues is -1 for every ket in (2.20.8). A measurement that distinguishes these four kets gives values to the top-row operators. For instance, when we detect $|++-\rangle$, the eigenvalues are $+1, +1, -1$ for the top row, while the operators in the middle and bottom rows do not have known values. There are quartets of such kets for each row and each column, see Exercise 49, and the corresponding measurements are *contexts* in which the three operators in one of the rows or one of the columns can have values.

In each of the six contexts, we could feel justified to regard the observed values — in that row or column — as pre-existing and revealed by the measurement, while we cannot say anything about the values of the other six operators. Perhaps, then, each of the nine operators does have a pre-

existing value and we are just limited in what we can reveal, namely at most three of the nine values?

No, this is not an option, clearly not. Try to assign $+1$ or -1 to each of the nine operators in Mermin's table (2.20.2), with due attention to the $+1$ products for the columns and the -1 products for the rows. You will surely fail because the column products require that the product of all nine entries equals $+1$, whereas the row products require -1. It follows that each row and each column can have values in a particular context and only then. A context-free assignment of values is not possible. Hence, the negative answer in (2.20.7).

In summary, quantum measurements do not reveal property values that have meaning before the measurement, unless you can predict the outcome with certainty — which you never can in a real-life laboratory experiment as there are always imperfections. The measurement result becomes known when the detection event happens and is taken note of, and these events are randomly realized. We can make statements about the probability of an event to happen, and that's all. This is the brutal fact of life that we recognized in Chapter 1.

2.21 Measurements with more than two outcomes

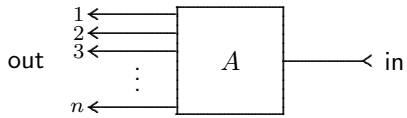
For a single silver atom, there are two outcomes in a Stern–Gerlach measurement, that is, the measurement probes a binary alternative, namely "spin up" or "spin down." This is an example of a *qubit*, the standard term for a quantum binary alternative or "quantum bit." The option of having superpositions of the alternatives distinguishes the qubit profoundly from the classical bit, which is exemplified by the binary alternative of coin tossing. Then, for a pair of atoms there are $2^2 = 4$ outcomes altogether for such measurements on both atoms, and so forth: $2^3 = 8$ for three atoms, $2^4 = 16$ for four atoms, ...; we have a pair, a trio, a quartet, ... of qubits, respectively.

Therefore, we need to extend the formalism such that there are 4, 8, 16, ... pairwise orthogonal ket vectors, rather than just two. Further, we note that instead of the magnetic properties of the atom, we could be interested in other physical phenomena, for which there are more than two choices to begin with. Think, for example, of the total electric charge of a collection of atoms. That could be 0, ± 1, ± 2, ± 3, ... units of electric charge, with no *a priori* restriction on these integers — an example of a physical property that would require an infinite set of pairwise orthogonal

kets in the mathematical formalism. And then there is the possibility of a continuum of measurement result, as one gets it when determining the position of an object, for example. Here, too, we need an infinite number of pairwise orthogonal kets, but a continuous set, rather than a discrete one as in the previous example of electric charge.

Further, in those situations where a finite number of kets suffice, it need not always be a power of 2; one can also have physical properties that have 3, 4, 5, 6, ... possible values as measurement results. For instance, there are atoms (orthohelium is an example), such that a Stern–Gerlach apparatus splits a beam of them into three.

It is clear, then, that we must generalize the formalism if we want, as we do, to deal with these other physical properties. For the time being, we shall assume that the number of possible measurement results is a finite number n and shall address the case of infinitely many measurements results later, in Chapter 4. So the symbolic generalization of a Stern–Gerlach apparatus is of the sort



(2.21.1)

We assign numbers $a_1, \ldots, a_n$ as measurement results to the various outcomes and have kets

$$\big|a_1\big\rangle\,,\ \big|a_2\big\rangle\,,\ \ldots\,,\ \big|a_n\big\rangle \tag{2.21.2}$$

associated with them. These exclude each other, just as $\uparrow_z$ and $\downarrow_z$ did in (2.8.4), which is expressed by their orthogonality

$$\big\langle a_j\big|a_k\big\rangle = \delta_{jk} = \begin{cases} 1 & \text{if } j = k\,, \\ 0 & \text{if } j \neq k\,, \end{cases} \tag{2.21.3}$$

for $j, k = 1, 2, \ldots, n$. Since this also contains the statement of normalization to unit length, $\langle a_j|a_j\rangle = 1$ for $j = 1, 2, \ldots, n$, the generalization of (2.8.3), we have an *orthonormality relation*. The symbol δ_{jk} is known as the *Kronecker* delta symbol*.

We did not overlook any possibilities, which is to say that the kets $|a_j\rangle$ are complete, as expressed by the *completeness relation*

*Leopold Kronecker (1823–1891)

$$\sum_{j=1}^{n} |a_j\rangle\langle a_j| = 1\,; \tag{2.21.4}$$

the kets $|a_j\rangle$ make up an orthogonal basis and so do the bras $\langle a_j| = |a_j\rangle^\dagger$. Applied to the arbitrary ket $|\,\rangle$,

$$|\,\rangle = \sum_{j=1}^{n} |a_j\rangle\langle a_j|\,\rangle\,, \tag{2.21.5}$$

the completeness relation states that $|\,\rangle$ is a weighted sum of the basis kets, with the probability amplitudes $\langle a_j|\,\rangle$ as weights. Their squares $\big|\langle a_j|\,\rangle\big|^2$ are then the probabilities that the measurement has outcome a_j if the ket $|\,\rangle$ correctly describes the incoming atoms. These probabilities sum up to unity,

$$\begin{aligned}\sum_j \big|\langle a_j|\,\rangle\big|^2 = \sum_j \langle\,|a_j\rangle\langle a_j|\,\rangle = \langle\,\Big|\Big(\sum_{j=1}^{n} |a_j\rangle\langle a_j|\Big)\Big|\,\rangle \\ = \langle\,|1|\,\rangle = \langle\,|\,\rangle = 1\,,\end{aligned} \tag{2.21.6}$$

which is just another way of saying that no possibility was left out.

Having assigned numbers a_j to the outcomes, we naturally use them as the eigenvalues of the operator A that we associate with the measured property,

$$A|a_j\rangle = |a_j\rangle a_j\,, \tag{2.21.7}$$

and the kets $|a_j\rangle$ are the respective eigenkets. The bras $\langle a_j|$ are the eigenbras, quite analogously,

$$\langle a_j|A = a_j\langle a_j|\,. \tag{2.21.8}$$

Together with the completeness relation in (2.21.4) and the orthonormality in (2.21.3), they require

$$A = \sum_{j=1}^{n} |a_j\rangle a_j\langle a_j| \tag{2.21.9}$$

for the operator A that we associate with the property measurement by the device in (2.21.1).

Note, however, that the two eigenvector equations (2.21.7) and (2.21.8) are not adjoints of each other. Rather the adjoint statement of (2.21.7) is

$$\langle a_j|A^\dagger = a_j^*\langle a_j| \tag{2.21.10}$$

consistent with

$$A^\dagger = \left(\sum_{j=1}^n |a_j\rangle a_j\langle a_j|\right)^\dagger = \sum_{j=1}^n |a_j\rangle a_j^*\langle a_j| \tag{2.21.11}$$

which is an immediate consequence of

$$\Big(|1\rangle\lambda\langle 2|\Big)^\dagger = |2\rangle\lambda^*\langle 1|\,, \tag{2.21.12}$$

a simple generalization of (2.8.18) that is used here for $|1\rangle = |a_j\rangle$, $\lambda = a_j$, $\langle 2| = \langle a_j|$, and $j = 1, \ldots, n$. In other words, the kets $|a_j\rangle$ are also eigenkets of $A^\dagger$ with eigenvalues a_j^*, and the bras $\langle a_j|$ are eigenbras of $A^\dagger$ with these eigenvalues.

Rather than the numbers a_j, we can assign another set of numbers to the n different possibilities, such as $f(a_j)$, where the function f must be well defined for all arguments $a_1, a_2, \ldots, a_n$ but we need not specify $f(a)$ for other values of a. Then, instead of A given above, we have the operator

$$\sum_{j=1}^n |a_j\rangle f(a_j)\langle a_j| \tag{2.21.13}$$

as the mathematical symbol for the physical property measured by the apparatus. It is systematic to regard the right-hand side as defining $f(A)$, the corresponding function of operator A,

$$f(A) = f\left(\sum_{j=1}^n |a_j\rangle a_j\langle a_j|\right) = \sum_{j=1}^n |a_j\rangle f(a_j)\langle a_j|\,. \tag{2.21.14}$$

The notation is suggestive, and as such it must not be misleading, that is to say that whenever we can give $f(A)$ meaning in a different way, the two definitions must coincide. So let us look at A^2,

$$\begin{aligned} A^2 &= \left(\sum_{j=1}^n |a_j\rangle a_j\langle a_j|\right)^2 \\ &= \sum_{j=1}^n\sum_{k=1}^n |a_j\rangle a_j \underbrace{\langle a_j|a_k\rangle}_{=\delta_{jk}} a_k\langle a_k| = \sum_{j=1}^n |a_j\rangle a_j^2\langle a_j|\,, \end{aligned} \tag{2.21.15}$$

and by induction we conclude that

$$A^N = \left(\sum_{j=1}^{n} |a_j\rangle a_j \langle a_j| \right)^N = \sum_{j=1}^{n} |a_j\rangle a_j^N \langle a_j| \,. \tag{2.21.16}$$

Accordingly,

$$f(A) = \sum_j |a_j\rangle f(a_j) \langle a_j| \tag{2.21.17}$$

is consistent if $f(A)$ is a power of A, and then it is also consistent when $f(A)$ is a polynomial of A. Since almost all functions of potential interest to us can be approximated by, or related to, polynomials, we accept (2.21.17) as the generally valid definition of a function of A. While the mapping $a \mapsto f(a)$ turns complex numbers into other complex numbers, the induced mapping $A \mapsto f(A)$ turns operators into other operators.

The right-hand side of (2.21.17) is called the *spectral decomposition* of $f(A)$, and we see the spectral decompositions of A itself and of $A^\dagger$ in (2.21.11). Thereby, we keep in mind that $f(x)$ must be well defined for $x = a_1$, $x = a_2$, ..., $x = a_n$, while other values of x are irrelevant. As a consequence of (2.21.17), two functions of A, $f_1(A)$ and $f_2(A)$, say, are equal if $f_1(a) = f_2(a)$ for all eigenvalues a of A and only in that situation. See Exercise 38 for an example.

A second physical property, B, has measurement results $b_1, b_2, \ldots, b_n$, eigenbras $\langle b_j|$, eigenkets $|b_k\rangle$, *et cetera* and for all relations stated above about A, there are, of course, the analogous statements about B, namely

$$\begin{aligned}
\text{eigenkets, eigenbras:} \quad & B|b_k\rangle = |b_k\rangle b_k \,, \quad \langle b_k|B = b_k \langle b_k| \,, \\
\text{orthonormality:} \quad & \langle b_j|b_k\rangle = \delta_{jk} \,, \\
\text{completeness:} \quad & \sum_{k=1}^{n} |b_k\rangle\langle b_k| = 1 \,, \\
\text{spectral decomposition:} \quad & B = \sum_{k=1}^{n} |b_k\rangle b_k \langle b_k| \,, \\
\text{function of } B\text{:} \quad & g(B) = \sum_{k=1}^{n} |b_k\rangle g(b_k) \langle b_k| \,.
\end{aligned} \tag{2.21.18}$$

And, in addition, we have statements involving both sets of kets and bras, in particular that of (2.8.19),

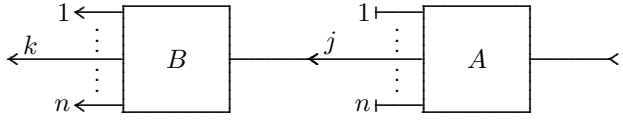
$$\langle a_j|b_k\rangle = \langle b_k|a_j\rangle^* \,, \tag{2.21.19}$$

that reversed probability amplitudes are complex conjugates of each other.

The fundamental symmetry of the probabilities

$$\text{prob}(b_k \leftarrow a_j) = \big|\langle b_k|a_j\rangle\big|^2 = \big|\langle a_j|b_k\rangle\big|^2 = \text{prob}(a_j \leftarrow b_k) \qquad (2.21.20)$$

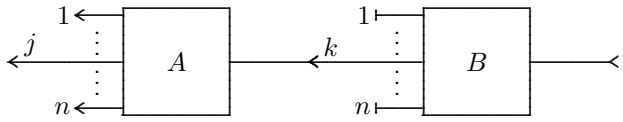
expresses an important fact about quantum processes. In the pictorial symbolization with the devices of (2.21.1), we would draw



measure property B and look for the kth outcome

measure property A, select atoms with the jth outcome (2.21.21)

for the experiment in which we measure $\text{prob}(b_k \leftarrow a_j)$ and



measure property A and look for the jth outcome

measure property B, select atoms with the kth outcome (2.21.22)

for the experiment in which we measure $\text{prob}(a_j \leftarrow b_k)$.

These are two quite different experiments, if properties A and B are quite different aspects of an atomic system, such as, say, its energy and its angular momentum. Nevertheless, and irrespective of which two properties we are dealing with, the equality of the probabilities (2.21.20) is always true. It seems, however, that this fundamental prediction of the quantum-mechanical formalism has never been studied systematically in an experiment, and so we must rely on circumstantial evidence provided by the overwhelmingly successful performance of quantum mechanics in much more complicated situations, where this symmetry is an essential ingredient in the calculations.

2.22 Unitary operators

Now, having the A states and the B states at hand, we can ask for the operator that relates the two ket bases to each other: What is U_{ba} in

$$U_{ba}|a_k\rangle = |b_k\rangle \qquad \text{for} \quad k = 1, 2, \ldots, n\,? \qquad (2.22.1)$$

Multiply by $\langle a_k|$ from the right, sum over k, and exploit the completeness relation for the A states,

$$\sum_k U_{ba}|a_k\rangle\langle a_k| = U_{ba}\underbrace{\sum_k |a_k\rangle\langle a_k|}_{=1} = \sum_k |b_k\rangle\langle a_k|\,, \tag{2.22.2}$$

to arrive at

$$U_{ba} = \sum_{k=1}^{n} |b_k\rangle\langle a_k|\,. \tag{2.22.3}$$

It follows that

$$\langle b_k|U_{ba} = \langle a_k| \qquad \text{for} \quad k = 1, 2, \ldots, n \tag{2.22.4}$$

holds as well.

The reverse mapping

$$U_{ab}|b_k\rangle = |a_k\rangle \tag{2.22.5}$$

is accomplished by the operator

$$U_{ab} = \sum_{k=1}^{n} |a_k\rangle\langle b_k|\,, \tag{2.22.6}$$

which is both the inverse of U_{ba},

$$U_{ba}U_{ab} = U_{ab}U_{ba} = 1: \quad U_{ab} = U_{ba}^{-1}\,, \tag{2.22.7}$$

and the adjoint of U_{ba} ,

$$\begin{aligned} U_{ab} &= \sum_k |a_k\rangle\langle b_k| = \sum_k \Big(|b_k\rangle\langle a_k|\Big)^\dagger \\ &= \left(\sum_k |b_k\rangle\langle a_k|\right)^\dagger = U_{ba}^\dagger\,. \end{aligned} \tag{2.22.8}$$

It follows that U_{ba} has the characteristic property that its adjoint is its inverse,

$$U_{ba}(U_{ba})^\dagger = 1\,, \qquad (U_{ba})^\dagger U_{ba} = 1\,, \tag{2.22.9}$$

and the same is, of course, also true for U_{ab},

$$U_{ab}(U_{ab})^\dagger = 1\,, \qquad (U_{ab})^\dagger U_{ab} = 1\,. \tag{2.22.10}$$

Operators of this kind are called *unitary*. They are the generalization of rotation operators for 3-vectors in ordinary space. For, a rotation leaves all scalar properties unchanged,

$$\left.\begin{array}{l} \boldsymbol{r} \xrightarrow[\text{rotation}]{} \boldsymbol{r}' \\ \boldsymbol{s} \xrightarrow[\text{rotation}]{} \boldsymbol{s}' \end{array}\right\} \quad \boldsymbol{r}\cdot\boldsymbol{s} = \boldsymbol{r}'\cdot\boldsymbol{s}' \tag{2.22.11}$$

for any pair of 3-vectors. The corresponding statement about unitary operators is the invariance of brackets, that is, of inner products,

$$\left.\begin{array}{l} \big|1\big\rangle \xrightarrow[\text{unitary}]{} \big|1'\big\rangle \\ \big|2\big\rangle \xrightarrow[\text{unitary}]{} \big|2'\big\rangle \end{array}\right\} \quad \big\langle 1\big|2\big\rangle = \big\langle 1'\big|2'\big\rangle \tag{2.22.12}$$

for all kets $|1\rangle, |2\rangle$.

Indeed, if we have

$$\begin{aligned} \big|1\big\rangle &\longrightarrow \big|1'\big\rangle = U\big|1\big\rangle\,, \\ \big|2\big\rangle &\longrightarrow \big|2'\big\rangle = U\big|2\big\rangle\,, \end{aligned} \tag{2.22.13}$$

then

$$\begin{aligned} \big\langle 1\big|2\big\rangle \longrightarrow \big\langle 1'\big|2'\big\rangle &= \Big(\big\langle 1\big|U^\dagger\big\rangle\Big)\Big(U\big|2\big\rangle\Big) \\ &= \big\langle 2\big|(U^\dagger U)\big|2\big\rangle \\ &= \big\langle 1\big|2\big\rangle \end{aligned} \tag{2.22.14}$$

holds for *any* choice of $|1\rangle$ and $|2\rangle$, if

$$U^\dagger U = 1\,, \tag{2.22.15}$$

that is, if U is unitary. In particular, it follows that a unitary operator maps any given set of orthonormal kets onto another orthonormal set. The unitary operators U_{ba} and U_{ab} from above illustrate this, inasmuch as they were constructed such that they map the bases composed of the $|a_k\rangle$s and the $|b_k\rangle$s onto each other.

Given some unitary operator, we denote its eigenvalues by u_k, the corresponding eigenkets by $|u_k\rangle$, and the eigenbras by $\langle u_k|$ so that

$$\begin{aligned} U\big|u_k\big\rangle &= \big|u_k\big\rangle u_k\,, \\ \big\langle u_k\big|U &= u_k\big\langle u_k\big|\,, \end{aligned} \tag{2.22.16}$$

which we supplement by the adjoint statements,

$$\begin{aligned} \langle u_k|U^\dagger &= u_k^*\langle u_k|\,, \\ U^\dagger|u_k\rangle &= |u_k\rangle u_k^*\,. \end{aligned} \tag{2.22.17}$$

We combine them in

$$\begin{aligned} u_k^*\langle u_k|u_k\rangle u_k &= \langle u_k|U^\dagger U|u_k\rangle \\ &= \langle u_k|u_k\rangle \end{aligned} \tag{2.22.18}$$

to conclude that

$$u_k^* u_k = 1 \qquad \text{or} \qquad |u_k| = 1 \qquad \text{for all} \quad k = 1, \ldots, n\,. \tag{2.22.19}$$

That is, all eigenvalues of a unitary operator are complex numbers with unit modulus; they are phase factors,

$$u_k = \mathrm{e}^{\mathrm{i}\varphi_k} \qquad \text{with a real phase } \varphi_k. \tag{2.22.20}$$

The converse is also true; see Exercise 56.

2.23 Hermitian operators

Given U, and thus its eigenvalues u_k, the phases φ_k are not determined uniquely; there is the obvious arbitrariness of adding any integer multiple of 2π to φ_k without changing the value of u_k. But, once we have chosen a particular set of φ_k, we can construct the new operator

$$\Phi = \sum_{k=1}^{n} |u_k\rangle \varphi_k \langle u_k|\,, \tag{2.23.1}$$

which is such that $U = \mathrm{e}^{\mathrm{i}\Phi}$. The fact that the phases φ_k are real — they are the eigenvalues of Φ, of course — implies that Φ is equal to its adjoint,

$$\begin{aligned} \Phi^\dagger &= \sum_{k=1}^{n} \Big(|u_k\rangle \varphi_k \langle u_k| \Big)^\dagger \\ &= \sum_{k=1}^{n} |u_k\rangle \varphi_k^* \langle u_k| = \sum_{k=1}^{n} |u_k\rangle \varphi_k \langle u_k| = \Phi\,. \end{aligned} \tag{2.23.2}$$

It is thus an example of another important class of operators, those that are adjoints of themselves. Such operators are called *selfadjoint* operators or *hermitian** operators; their eigenvalues are real.

The very important relation

$$U = \mathrm{e}^{\mathrm{i}H} \quad \text{with} \quad H = H^\dagger \quad \text{(hermitian)}$$
$$\text{and} \quad U = (U^\dagger)^{-1} \quad \text{(unitary)} \tag{2.23.3}$$

goes both ways: If U is unitary, then H is hermitian (see above), and if H is hermitian, then U is unitary. To demonstrate the latter, we use the spectral decomposition of H,

$$H = \sum_{k=1}^{n} \big|h_k\big\rangle h_k \big\langle h_k\big| \,, \tag{2.23.4}$$

where each h_k is a real eigenvalue and $|h_k\rangle$, $\langle h_k|$ are the corresponding eigenkets and eigenbras in

$$\mathrm{e}^{\mathrm{i}H} = \exp\left(\mathrm{i}\sum_{k=1}^{n} \big|h_k\big\rangle h_k \big\langle h_k\big|\right) = \sum_{k=1}^{n} \big|h_k\big\rangle\, \mathrm{e}^{\mathrm{i}h_k} \big\langle h_k\big| \,. \tag{2.23.5}$$

It follows that $\mathrm{e}^{\mathrm{i}H}$ is unitary because we have the situation of Exercise 56 with $A = H$ and $f(a) = \mathrm{e}^{\mathrm{i}a}$.

2.24 Hilbert spaces for kets and bras

In summary, quantum kinematics has kets and bras as the basic symbols for the description of quantum systems. A linear combination of kets, with complex coefficients, is another ket; likewise, a linear combination of bras is another bra: We have a linear space of kets and a linear space of bras. The two spaces are related by the adjoint operation,

$$\big|1\big\rangle^\dagger = \big\langle 1\big| \,, \quad \big\langle 2\big|^\dagger = \big|2\big\rangle \,, \tag{2.24.1}$$

that turns a ket into the corresponding bra and *vice versa*.

Bra-ket products are complex numbers. In view of the adjoint link between kets and bras, we can understand bra-kets as assigning numbers to pairs of kets or pairs of bras, and this assignment has all the properties of an inner product,

$$\big\langle 1\big|2\big\rangle = \Big(\big|1\big\rangle, \big|2\big\rangle\Big) = \Big(\big\langle 1\big|, \big\langle 2\big|\Big) \,, \tag{2.24.2}$$

*Charles Hermite (1822–1901)

as we first discussed in Section 2.8. Therefore, the ket space is a linear space endowed with an inner product, and the space has no gaps — in the way the rational numbers have gaps because they do not include lots of real numbers, such as $\sqrt{2}$. This is to say that the ket space is a Hilbert* space and the bra space is another Hilbert space. The adjoint operation is a one-to-one mapping between the two Hilbert spaces.

Then, we have the ket-bras, linear operators that map kets on kets and bras on bras. The unitary and the hermitian operators are particular and related to each other by exponentiation. Among the hermitian operators are the statistical operators that represent what we know about the physical systems.

The operators make up another linear space that becomes a Hilbert space when we endow it with an inner product in accordance with

$$(A, B) = \mathrm{tr}\big(A^\dagger B\big)\,, \tag{2.24.3}$$

the so-called Hilbert–Schmidt* inner product. We made use of this on various occasions such as (2.15.15) and Exercises 37 and 40.

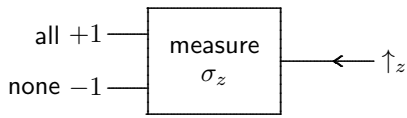
*David HILBERT (1862–1943) *Erhard SCHMIDT (1876–1959)

Chapter 3

Dynamics: How Quantum Systems Evolve

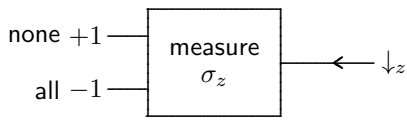
3.1 Schrödinger equation

Compare the following two situations that harken back to the end of Section 2.4. In the first scenario, atoms arrive "+ in z" and σ_z is measured, with outcome $+1$ for all of them:



(3.1.1)

or with outcome -1 for all if they arrive as $\downarrow_z$:

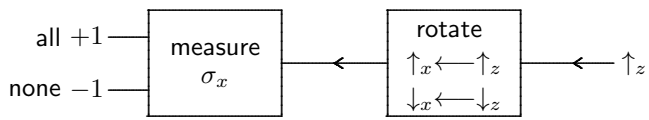


(3.1.2)

In the second scenario, a homogeneous magnetic field of appropriate strength and direction (along the y axis) effects

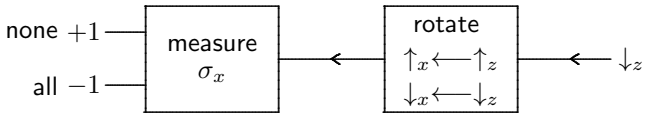
$$\uparrow_z \longrightarrow \uparrow_x \quad \text{and} \quad \downarrow_z \longrightarrow \downarrow_x , \tag{3.1.3}$$

and then σ_x is measured, so that



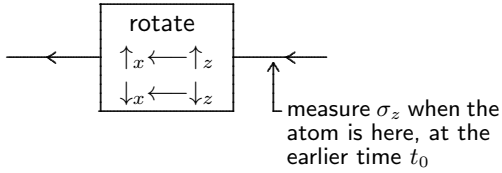
(3.1.4)

or



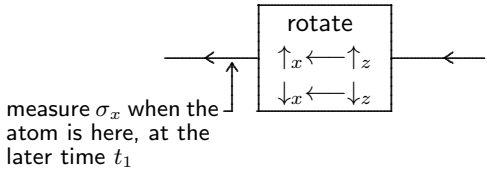
(3.1.5)

The measurement *result* is unchanged, you get predictably $+1$ for $\uparrow_z$ arriving and predictably -1 for $\downarrow_z$ arriving, although the apparatus is quite different. We can summarize these observations by stating that


rotate
$\uparrow_x \longleftarrow \uparrow_z$
$\downarrow_x \longleftarrow \downarrow_z$
measure σ_z when the atom is here, at the earlier time t_0

(3.1.6)

and


rotate
$\uparrow_x \longleftarrow \uparrow_z$
$\downarrow_x \longleftarrow \downarrow_z$
measure σ_x when the atom is here, at the later time t_1

(3.1.7)

are equivalent measurements as they give the same results. Now, recognizing that the instant at which you perform the measurement is crucial, we indicate the time dependence as an argument, such that $\sigma_z(t)$, for example, refers to measuring in the z direction at time t. In the above situation, then, we have

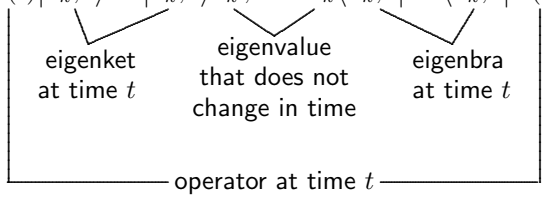
$$\sigma_x(t_1) = \sigma_z(t_0)\,. \tag{3.1.8}$$

This describes the overall effect, the "before to after" change, resulting from the action of the magnetic field, but clearly there must be intermediate stages with an altogether smooth transition from t_0 to t_1.

To formulate the equations of motion that describe infinitesimal changes during infinitesimal time intervals, we consider the general case where operator A has n different eigenvalues $a_1\,, \ldots\,, a_n$ and corresponding eigenkets and eigenbras. To begin with, we stick to the case illustrated by the $\sigma_z(t_0) = \sigma_x(t_1)$ example above, where the possible measurement results $(+1$

or -1) do not change in time. Here, then, we have the eigenket and eigenbra equations for operator $A(t)$,

$$A(t)\big|a_k,t\big\rangle = \big|a_k,t\big\rangle a_k\,, \qquad a_k\big\langle a_k,t\big| = \big\langle a_k,t\big|A(t)\,. \tag{3.1.9}$$



They are compactly summarized in the spectral decomposition of $A(t)$,

$$A(t) = \sum_{k=1}^{n} \big|a_k,t\big\rangle a_k \big\langle a_k,t\big|\,, \tag{3.1.10}$$

supplemented by the orthonormality relation,

$$\big\langle a_j,t\big|a_k,t\big\rangle = \delta_{jk}\,, \tag{3.1.11}$$

and the completeness relation,

$$\sum_{k=1}^{n} \big|a_k,t\big\rangle\big\langle a_k,t\big| = 1\,. \tag{3.1.12}$$

These statements are valid at all times t, whereby it is crucial that all kets and bras refer to the same time.

The time derivative of bra $\langle a_k,t|$,

$$\frac{\partial}{\partial t}\big\langle a_k,t\big| = \lim_{\tau\to 0}\frac{1}{\tau}\Big(\big\langle a_k,t+\tau\big| - \big\langle a_k,t\big|\Big)\,, \tag{3.1.13}$$

compares $\langle a_k,t|$ with $\langle a_k,t+\tau|$. Since we have an orthonormal, complete set of such bras at each time, they must be related to each other by an infinitesimal unitary transformation when τ is tiny,

$$\big\langle a_k,t+\tau\big| = \big\langle a_k,t\big|U_t(\tau) \tag{3.1.14}$$

with $U_t(\tau)^\dagger U_t(\tau) = 1$ and $U_t(\tau=0) = 1$. Upon writing

$$U_t(\tau) = 1 - \frac{\mathrm{i}}{\hbar}H(t)\tau + \underbrace{\cdots}_{\text{terms of order } \tau^2} \tag{3.1.15}$$

with a constant $\hbar$ of metrical dimension energy $\times$ time so that the operator $H(t)$ has the metrical dimension of energy, we have

$$U_t(\tau)^\dagger = 1 + \frac{\mathrm{i}}{\hbar} H(t)^\dagger \tau + \cdots \tag{3.1.16}$$

and get

$$U_t(\tau)^\dagger U_t(\tau) = 1 + \frac{\mathrm{i}}{\hbar}\big(H(t)^\dagger - H(t)\big)\tau + \cdots . \tag{3.1.17}$$

This tells us that $H(t)$ must be hermitian,

$$H(t)^\dagger = H(t) . \tag{3.1.18}$$

It is the hermitian generator of time transformations, the quantum analog of Hamilton's* function of classical mechanics. We borrow the terminology and call it the *Hamilton operator*. The classical Hamilton function is the energy of the evolving system; the quantum Hamilton operator is the energy operator. Its eigenvalues are the possible measurement results that one can get when measuring the energy, and the eigenkets and eigenbras of H describe states of the system with definite energy.

The constant $\hbar$, that is needed to render $U_t(\tau)$ dimensionless, is a natural constant — *Planck's*† *constant* divided by 2π. The numerical value is

$$\hbar = 1.05457 \times 10^{-34}\,\mathrm{J\,s} . \tag{3.1.19}$$

It is so very small because we use macroscopic units (Joule for energy, seconds for time) to express it. In atomic-scale units (electron-volts for energy, femto-seconds for time) the value of $\hbar$ is of order unity,

$$\hbar = 0.658212\,\mathrm{eV\,fs} , \tag{3.1.20}$$

as it should be.

With $U_t(\tau) = 1 - \frac{\mathrm{i}}{\hbar} H(t)\tau + \cdots$, we then have

$$\begin{aligned} \big\langle a_k, t+\tau\big| - \big\langle a_k, t\big| &= \big\langle a_k, t\big|\big(U_t(\tau) - 1\big) \\ &= -\frac{\mathrm{i}}{\hbar}\tau \big\langle a_k, t\big| H(t) + \cdots \end{aligned} \tag{3.1.21}$$

and

$$\mathrm{i}\hbar \frac{\partial}{\partial t}\big\langle a_k, t\big| = \big\langle a_k, t\big| H(t) \tag{3.1.22}$$

*William Rowan HAMILTON (1805–1865)

†Max Karl Ernst Ludwig PLANCK (1858–1947)

follows. There is nothing special about the operator $A(t)$, just as well the bra could refer to another operator, so we suppress the quantum number a_k and write more generally

$$\mathrm{i}\hbar\frac{\partial}{\partial t}\langle\ldots,t| = \langle\ldots,t|H(t)\,, \tag{3.1.23}$$

where the ellipses means the same quantum numbers on both sides. This equation of motion is the celebrated *Schrödinger equation* — actually the one for bras, which is equivalent to, but not identical with, the original form of the equation that Schrödinger discovered. The adjoint of (3.1.23),

$$-\mathrm{i}\hbar\frac{\partial}{\partial t}|\ldots,t\rangle = H(t)|\ldots,t\rangle, \tag{3.1.24}$$

is the Schrödinger equation for kets, and there is a plethora of numerical versions of the Schrödinger equation.

3.2 Heisenberg equation

The time derivative of operator $A(t)$,

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}A(t) &= \sum_k\left(\frac{\partial|a_k,t\rangle}{\partial t}a_k\langle a_k,t| + |a_k,t\rangle a_k\frac{\partial\langle a_k,t|}{\partial t}\right)\\ &= \sum_k\left(-\frac{1}{\mathrm{i}\hbar}H(t)|a_k,t\rangle a_k\langle a_k,t| + |a_k,t\rangle a_k\langle a_k,t|\frac{1}{\mathrm{i}\hbar}H(t)\right)\\ &= -\frac{1}{\mathrm{i}\hbar}H(t)A(t) + \frac{1}{\mathrm{i}\hbar}A(t)H(t)\,,\end{aligned} \tag{3.2.1}$$

is conveniently written as

$$\frac{\mathrm{d}}{\mathrm{d}t}A(t) = \frac{1}{\mathrm{i}\hbar}[A(t),H(t)]\,, \tag{3.2.2}$$

where

$$[X,Y] = XY - YX = -[Y,X] \tag{3.2.3}$$

denotes the so-called *commutator* of operators X and Y. It has a lot in common with a differentiation. In particular, there is a sum rule

$$\begin{aligned}[X_1+X_2,Y] &= [X_1,Y]+[X_2,Y]\,,\\ [X,Y_1+Y_2] &= [X,Y_1]+[X,Y_2]\,,\end{aligned} \tag{3.2.4}$$

and a product rule,

$$[X_1X_2, Y] = [X_1, Y]X_2 + X_1[X_2, Y]\,,$$
$$[X, Y_1Y_2] = [X, Y_1]Y_2 + Y_1[X, Y_2]\,, \tag{3.2.5}$$

where due attention must be paid to the order in which the operators appear in these products. These relations, and many more, are easily verified by inspection. In particular, there is a Jacobi* identity for commutators,

$$\big[[X, Y], Z\big] + \big[[Y, Z], X\big] + \big[[Z, X], Y\big] = 0\,; \tag{3.2.6}$$

see Exercise 60.

A bit more general than $A(t)$ with time-independent eigenvalues is the case where this restriction is lifted. As a consequence of such a *parametric* time dependence, we then have an extra term

$$\sum_k \big|a_k, t\big\rangle \frac{\partial a_k}{\partial t} \big\langle a_k, t\big| = \frac{\partial A}{\partial t} \tag{3.2.7}$$

in

$$\frac{\mathrm{d}}{\mathrm{d}t} A = \frac{\partial A}{\partial t} + \frac{1}{\mathrm{i}\hbar}[A, H]\,. \tag{3.2.8}$$

This is *Heisenberg's equation of motion.* By contrast, the commutator term $\frac{1}{\mathrm{i}\hbar}[A, H]$ originates in physical interactions, as they are contributing various amounts of interaction energy to the Hamilton operator H, so that $\frac{1}{\mathrm{i}\hbar}[A, H]$ refers to the *dynamical* time dependence of A.

As a rule, a parametric time dependence results from human interference. For instance, we choose to assign different numbers to the measurement results at different times because for some reason that is convenient. Or, we actually change a parameter of the apparatus in question, such as altering the direction and strength of the magnetic field that rotates the magnetic moments in the setups in (3.1.4) and (3.1.7). The magnetic interaction energy, $-\boldsymbol{\mu} \cdot \boldsymbol{B}(t)$, then introduces the parametric time dependence into the Hamilton operator. It is, in fact, the only net time dependence that H can possess because $[H, H] = 0$ holds always, so that Heisenberg's equation says

$$\frac{\mathrm{d}}{\mathrm{d}t} H = \frac{\partial}{\partial t} H \tag{3.2.9}$$

for the Hamilton operator itself: The time dependence of H is always parametric in nature.

*Carl Gustav Jacob JACOBI (1804–1851)

3.3 Equivalent Hamilton operators

In the Heisenberg equation of motion (3.2.8) for A, the Hamilton operator H enters only in the commutator and, therefore, nothing changes if we add a constant ($\equiv$ a multiple of the identity operator) to H,

$$H \to H + \hbar\Omega : \quad \frac{\mathrm{d}}{\mathrm{d}t}A(t) \quad \text{unchanged.} \tag{3.3.1}$$

But, there is a change in the Schrödinger equation, so that we must carefully compare

$$\mathrm{i}\hbar\frac{\partial}{\partial t}\langle \ldots, t| = \langle \ldots, t|H \qquad \text{(original)} \tag{3.3.2}$$

with

$$\mathrm{i}\hbar\frac{\partial}{\partial t}\langle \ldots, t|' = \langle \ldots, t|'(H + \hbar\Omega) \qquad \text{(altered)}, \tag{3.3.3}$$

where the change of the Hamilton operator is accompanied by a corresponding change of the bras,

$$\langle \ldots, t| \to \langle \ldots, t|' \,. \tag{3.3.4}$$

We have

$$\left(\mathrm{i}\hbar\frac{\partial}{\partial t} - \hbar\Omega\right)\langle \ldots, t|' = \langle \ldots, t|'H \tag{3.3.5}$$

and note that the differential operator on the left can be written as

$$\mathrm{i}\hbar\frac{\partial}{\partial t} - \hbar\Omega = \mathrm{i}\hbar\left(\frac{\partial}{\partial t} + \mathrm{i}\Omega\right) = \mathrm{i}\hbar\,\mathrm{e}^{-\mathrm{i}\Omega t}\frac{\partial}{\partial t}\,\mathrm{e}^{\mathrm{i}\Omega t} \,. \tag{3.3.6}$$

(Keep in mind that there is an unwritten function of t to the right of these expressions, the function that is differentiated.) Thus, equivalently, we have

$$\mathrm{i}\hbar\frac{\partial}{\partial t}\Big(\mathrm{e}^{\mathrm{i}\Omega t}\langle \ldots, t|'\Big) = \Big(\mathrm{e}^{\mathrm{i}\Omega t}\langle \ldots, t|'\Big)H \tag{3.3.7}$$

and the identification

$$\langle \ldots, t|' = \mathrm{e}^{-\mathrm{i}\Omega t}\langle \ldots, t| \tag{3.3.8}$$

turns this into the original Schrödinger equation for $\langle \ldots, t|$.

The phase factor $\mathrm{e}^{-\mathrm{i}\Omega t}$ that is acquired by the bra vectors is the *same* for all bras, there is no dependence on the implicit quantum numbers indicated by the ellipsis. As a consequence, all bras are just multiplied by a

common phase factor, a freedom we have anyway. Such a phase factor is of no physical relevance, although it might be required for mathematical consistency. Since the kets acquire the complex conjugate phase factor

$$\big|\ldots,t\big\rangle \to \big|\ldots,t\big\rangle' = \big|\ldots,t\big\rangle\, \mathrm{e}^{\mathrm{i}\Omega t}\,, \tag{3.3.9}$$

no visible phase factors appear in the spectral decomposition of (3.1.10),

$$A = \sum_{k=1}^{n} \big|a_k,t\big\rangle a_k \big\langle a_k,t\big| = \sum_{k=1}^{n} \big|a_k,t\big\rangle' a_k \big\langle a_k,t\big|'\,. \tag{3.3.10}$$

Clearly, there is no physical significance of the replacement $H \to H + \hbar\Omega$. We can redefine the origin of the energy scale anyway we wish: *energy differences* are physically significant, not absolute energy values. Exercise 63 deals with the possibility of a time-dependent energy shift.

3.4 Von Neumann equation

We noted above that the Hamilton operator is particular because it has no dynamical time dependence, only a parametric one — very often none at all, when the external parameters do not change in time. Another special operator is the statistical operator ρ that summarizes our knowledge about the preparation of the system. As such, it is the mathematical symbol for phrases like "atoms that are $\uparrow_z$ at time t_0" which, by their very nature, refer to one particular moment in time, namely the instant at which we learned something about the atoms. As a consequence, a statistical operator has no total time dependence, $\dfrac{\mathrm{d}}{\mathrm{d}t}\rho = 0$, it is simply constant in time, as an operator. The Heisenberg equation of motion,

$$0 = \frac{\mathrm{d}}{\mathrm{d}t}\rho = \frac{\partial}{\partial t}\rho + \frac{1}{\mathrm{i}\hbar}[\rho, H]\,, \tag{3.4.1}$$

then says that the parametric time dependence of ρ — term $\dfrac{\partial}{\partial t}\rho$ — and the dynamical time dependence — term $\dfrac{1}{\mathrm{i}\hbar}[\rho, H]$ — compensate for each other exactly. This equation of motion for the statistical operator is often referred to as the *von Neumann* equation*; we regard it as a special case of the Heisenberg equation of motion.

*John (János) VON NEUMANN (1903–1957)

3.5 Example: Larmor precession

As an example, let us consider the statistical operator for magnetic silver atoms of (2.15.15),

$$\rho = \frac{1}{2}\big(1 + \boldsymbol{s}(t)\cdot\boldsymbol{\sigma}(t)\big)\,, \tag{3.5.1}$$

parametric time dependence — dynamical time dependence

where we now make explicit that both the Pauli vector operator $\boldsymbol{\sigma}$ and its expectation value $\boldsymbol{s} = \langle\boldsymbol{\sigma}\rangle = \mathrm{tr}(\rho\,\boldsymbol{\sigma})$ depend on time. The time dependence of $\boldsymbol{\sigma}(t)$ is dynamical, whereas that of $\boldsymbol{s}(t)$ is parametric: A change of $\boldsymbol{s}$ changes the functional dependence of ρ on the dynamical variables $\sigma_x, \sigma_y, \sigma_z$.

Here, the two time dependences compensate for each other, so that the product $\boldsymbol{s}(t)\cdot\boldsymbol{\sigma}(t)$ as a whole is time-independent. In particular, then, the eigenvalues of ρ do not change in time, and since they are

$$\frac{1}{2} \pm \frac{1}{2}\big|\boldsymbol{s}(t)\big|\,, \tag{3.5.2}$$

the length of the numerical vector $\boldsymbol{s}(t)$ is constant in time. Therefore, the only change of $\boldsymbol{s}$ from t to $t + \mathrm{d}t$ can be an infinitesimal rotation,

$$\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{s}(t) = \boldsymbol{\omega}(t)\times\boldsymbol{s}(t)\,, \tag{3.5.3}$$

where the angular velocity vector $\boldsymbol{\omega}(t)$ could itself depend on time. Since $\boldsymbol{s}(t) = \big\langle\boldsymbol{\sigma}(t)\big\rangle$, we infer that

$$\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{\sigma}(t) = \boldsymbol{\omega}(t)\times\boldsymbol{\sigma}(t)\,. \tag{3.5.4}$$

Indeed, then

$$\frac{\mathrm{d}}{\mathrm{d}t}(\boldsymbol{s}\cdot\boldsymbol{\sigma}) = \frac{\mathrm{d}\boldsymbol{s}}{\mathrm{d}t}\cdot\boldsymbol{\sigma} + \boldsymbol{s}\cdot\frac{\mathrm{d}\boldsymbol{\sigma}}{\mathrm{d}t} = (\boldsymbol{\omega}\times\boldsymbol{s})\cdot\boldsymbol{\sigma} + \boldsymbol{s}\cdot(\boldsymbol{\omega}\times\boldsymbol{\sigma}) = 0\,. \tag{3.5.5}$$

The right-hand side of (3.5.4) is the right-hand side of Heisenberg's equation $\dfrac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{\sigma} = \dfrac{1}{\mathrm{i}\hbar}[\boldsymbol{\sigma}, H]$, so that

$$\big[\boldsymbol{\sigma}, H\big] = \mathrm{i}\hbar\boldsymbol{\omega}\times\boldsymbol{\sigma}\,. \tag{3.5.6}$$

The Hamilton operator must be a function of $\boldsymbol{\sigma}$, there are no other operators available to construct it, and since all functions of $\boldsymbol{\sigma}$ are linear functions,

we have

$$H = \hbar \boldsymbol{\Omega} \cdot \boldsymbol{\sigma} \tag{3.5.7}$$

with $\boldsymbol{\Omega}(t)$ related to $\boldsymbol{\omega}(t)$. There is the further possibility of adding a multiple of the identity to H but, as discussed in Section 3.3, this is of no consequence, so we do not have to take it into account.

To find the relation between $\boldsymbol{\omega}$ and $\boldsymbol{\Omega}$, we recall (2.9.13),

$$\begin{aligned} \boldsymbol{a} \cdot \boldsymbol{\sigma}\ \boldsymbol{b} \cdot \boldsymbol{\sigma} &= \boldsymbol{a} \cdot \boldsymbol{b} + \mathrm{i}(\boldsymbol{a} \times \boldsymbol{b}) \cdot \boldsymbol{\sigma}\,, \\ \boldsymbol{b} \cdot \boldsymbol{\sigma}\ \boldsymbol{a} \cdot \boldsymbol{\sigma} &= \boldsymbol{a} \cdot \boldsymbol{b} - \mathrm{i}(\boldsymbol{a} \times \boldsymbol{b}) \cdot \boldsymbol{\sigma}\,, \end{aligned} \tag{3.5.8}$$

which are valid for all numerical 3-vectors $\boldsymbol{a}$ and $\boldsymbol{b}$. Their difference is the commutation relation

$$\big[\boldsymbol{a} \cdot \boldsymbol{\sigma}, \boldsymbol{b} \cdot \boldsymbol{\sigma}\big] = 2\mathrm{i}(\boldsymbol{a} \times \boldsymbol{b}) \cdot \boldsymbol{\sigma} = 2\mathrm{i}\boldsymbol{a} \cdot (\boldsymbol{b} \times \boldsymbol{\sigma})\,, \tag{3.5.9}$$

so that

$$\big[\boldsymbol{\sigma}, \boldsymbol{b} \cdot \boldsymbol{\sigma}\big] = 2\mathrm{i}\boldsymbol{b} \times \boldsymbol{\sigma}\,, \tag{3.5.10}$$

after getting rid of the arbitrary vector $\boldsymbol{a}$. The comparison with

$$\begin{aligned} [\boldsymbol{\sigma}, H] = [\boldsymbol{\sigma}, \hbar\boldsymbol{\Omega} \cdot \boldsymbol{\sigma}] &= 2\mathrm{i}\hbar\boldsymbol{\Omega} \times \boldsymbol{\sigma} \\ &= \mathrm{i}\hbar\boldsymbol{\omega} \times \boldsymbol{\sigma} \end{aligned} \tag{3.5.11}$$

establishes $\boldsymbol{\Omega} = \frac{1}{2}\boldsymbol{\omega}$, and we arrive at

$$H = \frac{1}{2}\hbar\boldsymbol{\omega} \cdot \boldsymbol{\sigma}\,. \tag{3.5.12}$$

Now, to be more specific, let us consider $\boldsymbol{\omega} = \omega\boldsymbol{e}_z$ with $\omega > 0$ independent of time t. Then

$$\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{\sigma} = \omega\boldsymbol{e}_z \times \boldsymbol{\sigma}\,, \tag{3.5.13}$$

with the components

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}\sigma_x &= \boldsymbol{e}_x \cdot (\omega\boldsymbol{e}_z \times \boldsymbol{\sigma}) = \omega \underbrace{(\boldsymbol{e}_x \times \boldsymbol{e}_z)}_{=-\boldsymbol{e}_y} \cdot \boldsymbol{\sigma} = -\omega\sigma_y\,, \\ \frac{\mathrm{d}}{\mathrm{d}t}\sigma_y &= \boldsymbol{e}_y \cdot (\omega\boldsymbol{e}_z \times \boldsymbol{\sigma}) = \omega \underbrace{(\boldsymbol{e}_y \times \boldsymbol{e}_z)}_{=\boldsymbol{e}_x} \cdot \boldsymbol{\sigma} = \omega\sigma_x\,, \\ \frac{\mathrm{d}}{\mathrm{d}t}\sigma_z &= \boldsymbol{e}_z \cdot (\omega\boldsymbol{e}_z \times \boldsymbol{\sigma}) = 0\,, \end{aligned} \tag{3.5.14}$$

which are, of course, exactly what we get from Heisenberg's equations directly,

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\sigma_x &= \frac{1}{\mathrm{i}\hbar}[\sigma_x, H] = \frac{1}{2\mathrm{i}}\omega \underbrace{[\sigma_x, \sigma_z]}_{=-2\mathrm{i}\sigma_y} = -\omega\sigma_y\,, \\
\frac{\mathrm{d}}{\mathrm{d}t}\sigma_y &= \frac{1}{\mathrm{i}\hbar}[\sigma_y, H] = \frac{1}{2\mathrm{i}}\omega \underbrace{[\sigma_y, \sigma_z]}_{=2\mathrm{i}\sigma_x} = \omega\sigma_x\,, \\
\frac{\mathrm{d}}{\mathrm{d}t}\sigma_z &= \frac{1}{\mathrm{i}\hbar}[\sigma_z, H] = \frac{1}{2\mathrm{i}}\omega \underbrace{[\sigma_z, \sigma_z]}_{=0} = 0\,,
\end{aligned} \tag{3.5.15}$$

where the Hamilton operator

$$H = \frac{1}{2}\hbar\omega\sigma_z \tag{3.5.16}$$

is obtained from (3.5.12) for $\boldsymbol{\omega} = \omega\boldsymbol{e}_z$.

We have to solve these equations in order to express $\boldsymbol{\sigma}(t)$ in terms of $\boldsymbol{\sigma}(0)$. The solution for σ_z is immediate,

$$\sigma_z(t) = \sigma_z(0)\,, \tag{3.5.17}$$

and to find that for σ_x and σ_y, we look at $\sigma_x + \mathrm{i}\sigma_y$,

$$\frac{\mathrm{d}}{\mathrm{d}t}(\sigma_x + \mathrm{i}\sigma_y) = -\omega\sigma_y + \mathrm{i}\omega\sigma_x = \mathrm{i}\omega(\sigma_x + \mathrm{i}\sigma_y)\,, \tag{3.5.18}$$

which is solved by

$$\sigma_x(t) + \mathrm{i}\sigma_y(t) = \mathrm{e}^{\mathrm{i}\omega t}\big(\sigma_x(0) + \mathrm{i}\sigma_y(0)\big)\,. \tag{3.5.19}$$

By taking the adjoint, we also get

$$\sigma_x(t) - \mathrm{i}\sigma_y(t) = \mathrm{e}^{-\mathrm{i}\omega t}\big(\sigma_x(0) - \mathrm{i}\sigma_y(0)\big)\,, \tag{3.5.20}$$

and then

$$\begin{aligned}
\sigma_x(t) &= \sigma_x(0)\cos(\omega t) - \sigma_y(0)\sin(\omega t)\,, \\
\sigma_y(t) &= \sigma_y(0)\cos(\omega t) + \sigma_x(0)\sin(\omega t)
\end{aligned} \tag{3.5.21}$$

by adding and subtracting. It is a matter of inspection to verify that they obey the differential equations (3.5.15) and reduce to an identity for $t = 0$.

Since $\boldsymbol{s} = \langle\boldsymbol{\sigma}\rangle$, we also have

$$\begin{aligned}
s_x(t) &= s_x(0)\cos(\omega t) - s_y(0)\sin(\omega t)\,, \\
s_y(t) &= s_y(0)\cos(\omega t) + s_x(0)\sin(\omega t)\,,
\end{aligned} \tag{3.5.22}$$

and

$$s_z(t) = s_z(0)\,, \tag{3.5.23}$$

showing once more, now rather explicitly, that the 3-vector $\boldsymbol{s}$ precesses around the axis specified by $\boldsymbol{\omega}$, here the z axis, with angular velocity $\omega = |\boldsymbol{\omega}|$. It takes the time $2\pi/\omega$ to complete one revolution, so we have a periodic motion with this period. In the context of Exercise 64, this is the so-called *Larmor* precession* of a magnetic moment in a homogeneous magnetic field.

3.6 Time-dependent probability amplitudes

With the Hamilton operator (3.5.16), the Schrödinger equations for eigenbras of σ_z are

$$\mathrm{i}\hbar\frac{\partial}{\partial t}\langle\uparrow_z, t| = \langle\uparrow_z, t|H(t) = \langle\uparrow_z, t|\frac{\hbar}{2}\omega\sigma_z(t) = \frac{\hbar}{2}\omega\langle\uparrow_z, t| \tag{3.6.1}$$

and, likewise,

$$\mathrm{i}\hbar\frac{\partial}{\partial t}\langle\downarrow_z, t| = -\frac{\hbar}{2}\omega\langle\downarrow_z, t|\,, \tag{3.6.2}$$

which are solved by

$$\begin{aligned}\langle\uparrow_z, t| &= \mathrm{e}^{-\frac{\mathrm{i}}{2}\omega t}\langle\uparrow_z, 0|\,,\\ \langle\downarrow_z, t| &= \mathrm{e}^{\frac{\mathrm{i}}{2}\omega t}\langle\downarrow_z, 0|\,.\end{aligned} \tag{3.6.3}$$

We use them to find the time-dependent probability amplitudes for a given state ket,

$$\begin{aligned}|\ \rangle &= |\uparrow_z, 0\rangle\alpha_0 + |\downarrow_z, 0\rangle\beta_0\\ &= |\uparrow_z, t\rangle\alpha(t) + |\downarrow_z, t\rangle\beta(t)\,,\end{aligned} \tag{3.6.4}$$

where

$$\alpha(t) = \langle\uparrow_z, t|\ \rangle = \mathrm{e}^{-\frac{\mathrm{i}}{2}\omega t}\langle\uparrow_z, 0|\ \rangle = \mathrm{e}^{-\frac{\mathrm{i}}{2}\omega t}\alpha_0 \tag{3.6.5}$$

and

$$\beta(t) = \langle\downarrow_z, t|\ \rangle = \mathrm{e}^{\frac{\mathrm{i}}{2}\omega t}\langle\downarrow_z, 0|\ \rangle = \mathrm{e}^{\frac{\mathrm{i}}{2}\omega t}\beta_0\,. \tag{3.6.6}$$

*Joseph LARMOR (1857–1942)

As a check of consistency, let us verify that they give us the same expectation values of $\sigma_x, \sigma_y, \sigma_z$ as a function of t. We recall the relations of Exercise 30,

$$\langle\sigma_x\rangle = 2\,\mathrm{Re}(\alpha^*\beta)\,, \quad \langle\sigma_y\rangle = 2\,\mathrm{Im}(\alpha^*\beta)\,, \quad \langle\sigma_z\rangle = |\alpha|^2 - |\beta|^2\,, \tag{3.6.7}$$

and thus find

$$\big\langle\sigma_x(t)\big\rangle = 2\,\mathrm{Re}\Big(\,\mathrm{e}^{\mathrm{i}\omega t}\alpha_0^*\beta_0\Big)\,. \tag{3.6.8}$$

With Euler's identity (2.5.48) and the initial values

$$\alpha_0^*\beta_0 = \frac{1}{2}\big\langle\sigma_x(0)\big\rangle + \frac{\mathrm{i}}{2}\big\langle\sigma_y(0)\big\rangle\,, \tag{3.6.9}$$

this becomes

$$\begin{aligned}\big\langle\sigma_x(t)\big\rangle &= \mathrm{Re}\Big(\big(\cos(\omega t) + \mathrm{i}\sin(\omega t)\big)\Big(\big\langle\sigma_x(0)\big\rangle + \mathrm{i}\,\big\langle\sigma_y(0)\big\rangle\Big)\Big)\\ &= \big\langle\sigma_x(0)\big\rangle\cos(\omega t) - \big\langle\sigma_y(0)\big\rangle\sin(\omega t)\,,\end{aligned} \tag{3.6.10}$$

and likewise we get

$$\begin{aligned}\big\langle\sigma_y(t)\big\rangle &= \mathrm{Im}\Big(\big(\cos(\omega t) + \mathrm{i}\sin(\omega t)\big)\Big(\big\langle\sigma_x(0)\big\rangle + \mathrm{i}\,\big\langle\sigma_y(0)\big\rangle\Big)\Big)\\ &= \big\langle\sigma_y(0)\big\rangle\cos(\omega t) + \big\langle\sigma_x(0)\big\rangle\sin(\omega t)\,,\end{aligned} \tag{3.6.11}$$

and finally,

$$\begin{aligned}\big\langle\sigma_z(t)\big\rangle &= \Big|\,\mathrm{e}^{-\frac{\mathrm{i}}{2}\omega t}\alpha_0\Big|^2 - \Big|\,\mathrm{e}^{\frac{\mathrm{i}}{2}\omega t}\beta_0\Big|^2\\ &= |\alpha_0|^2 - |\beta_0|^2 = \big\langle\sigma_z(0)\big\rangle\,.\end{aligned} \tag{3.6.12}$$

These are, of course, just the equations for the components of $\boldsymbol{s}(t)$ in (3.5.22) and (3.5.23).

3.7 Schrödinger equation for probability amplitudes

Lifting the restriction of $\boldsymbol{\omega} = \omega\boldsymbol{e}_z$, we return to the more general case (3.5.12) of an arbitrary, perhaps time-dependent, angular velocity vector $\boldsymbol{\omega}$,

$$H = \frac{1}{2}\hbar\boldsymbol{\omega}\cdot\boldsymbol{\sigma}\,. \tag{3.7.1}$$

The standard matrix representation of H, which refers to eigenkets and eigenbras of $\sigma_z(t)$, reads

$$H = \frac{\hbar}{2}\boldsymbol{\omega}\cdot\boldsymbol{\sigma}(t) = \Big(\,|{\uparrow_z},t\rangle\;\;|{\downarrow_z},t\rangle\,\Big)\mathcal{H}\begin{pmatrix}\langle{\uparrow_z},t|\\ \langle{\downarrow_z},t|\end{pmatrix} \tag{3.7.2}$$

with

$$\begin{aligned}\mathcal{H} &= \begin{pmatrix}\langle{\uparrow_z},t|\\ \langle{\downarrow_z},t|\end{pmatrix} H \Big(\,|{\uparrow_z},t\rangle\;\;|{\downarrow_z},t\rangle\,\Big)\\ &= \begin{pmatrix}\langle{\uparrow_z},t|H|{\uparrow_z},t\rangle & \langle{\uparrow_z},t|H|{\downarrow_z},t\rangle\\ \langle{\downarrow_z},t|H|{\uparrow_z},t\rangle & \langle{\downarrow_z},t|H|{\downarrow_z},t\rangle\end{pmatrix}\\ &= \frac{\hbar}{2}\begin{pmatrix}\omega_z & \omega_x-\mathrm{i}\omega_y\\ \omega_x+\mathrm{i}\omega_y & -\omega_z\end{pmatrix}.\end{aligned} \tag{3.7.3}$$

This matrix appears in the Schrödinger equation obeyed by the column of probability amplitudes, as specified by

$$|\;\rangle = |{\uparrow_z},t\rangle\alpha(t) + |{\downarrow_z},t\rangle\beta(t) \mathrel{\widehat{=}} \begin{pmatrix}\alpha(t)\\ \beta(t)\end{pmatrix}, \tag{3.7.4}$$

inasmuch as

$$\begin{aligned}\mathrm{i}\hbar\frac{\partial}{\partial t}\begin{pmatrix}\alpha(t)\\ \beta(t)\end{pmatrix} &= \mathrm{i}\hbar\frac{\partial}{\partial t}\begin{pmatrix}\langle{\uparrow_z},t|\\ \langle{\downarrow_z},t|\end{pmatrix}|\;\rangle\\ &= \begin{pmatrix}\langle{\uparrow_z},t|\\ \langle{\downarrow_z},t|\end{pmatrix}H|\;\rangle\\ &= \mathcal{H}\begin{pmatrix}\langle{\uparrow_z},t|\\ \langle{\downarrow_z},t|\end{pmatrix}|\;\rangle = \mathcal{H}\begin{pmatrix}\alpha(t)\\ \beta(t)\end{pmatrix},\end{aligned} \tag{3.7.5}$$

that is,

$$\mathrm{i}\hbar\frac{\partial}{\partial t}\begin{pmatrix}\alpha\\ \beta\end{pmatrix} = \mathcal{H}\begin{pmatrix}\alpha\\ \beta\end{pmatrix}. \tag{3.7.6}$$

Please note that in relations such as

$$\begin{pmatrix}\langle{\uparrow_z},t|\\ \langle{\downarrow_z},t|\end{pmatrix}H = \mathcal{H}\begin{pmatrix}\langle{\uparrow_z},t|\\ \langle{\downarrow_z},t|\end{pmatrix}, \tag{3.7.7}$$

the operator H acts from the right on the bras, whereas its numerical 2×2 matrix $\mathcal{H}$ multiplies the column from the left. In other words, both the

operator and the matrix act from their natural sides on the column of the bras.

As an example, consider

$$H = \frac{\hbar}{2}\omega\sigma_x \mathrel{\widehat{=}} \mathcal{H} = \frac{\hbar}{2}\omega\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \tag{3.7.8}$$

for which

$$\mathrm{i}\hbar\frac{\partial}{\partial t}\begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \frac{\hbar}{2}\omega\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \frac{\hbar}{2}\omega\begin{pmatrix} \beta \\ \alpha \end{pmatrix} \tag{3.7.9}$$

or

$$\frac{\partial}{\partial t}\alpha = -\frac{\mathrm{i}}{2}\omega\beta, \qquad \frac{\partial}{\partial t}\beta = -\frac{\mathrm{i}}{2}\omega\alpha\,. \tag{3.7.10}$$

These simplify if we take the sum and difference of α and β,

$$\begin{aligned} \frac{\partial}{\partial t}(\alpha+\beta) &= -\frac{\mathrm{i}\omega}{2}(\alpha+\beta)\,, \\ \frac{\partial}{\partial t}(\alpha-\beta) &= \frac{\mathrm{i}\omega}{2}(\alpha-\beta)\,. \end{aligned} \tag{3.7.11}$$

Their solutions,

$$\begin{aligned} \alpha(t)+\beta(t) &= \mathrm{e}^{-\mathrm{i}\omega t/2}(\alpha_0+\beta_0)\,, \\ \alpha(t)-\beta(t) &= \mathrm{e}^{\mathrm{i}\omega t/2}(\alpha_0-\beta_0)\,, \end{aligned} \tag{3.7.12}$$

tell us that

$$\begin{aligned} \alpha(t) &= \alpha_0\cos\left(\frac{\omega t}{2}\right) - \mathrm{i}\beta_0\sin\left(\frac{\omega t}{2}\right), \\ \beta(t) &= \beta_0\cos\left(\frac{\omega t}{2}\right) - \mathrm{i}\alpha_0\sin\left(\frac{\omega t}{2}\right). \end{aligned} \tag{3.7.13}$$

One verifies by inspection that they obey the coupled differential equations of (3.7.10) and reduce to an identity for $t = 0$ when $\alpha(0) = \alpha_0$ and $\beta(0) = \beta_0$.

This pattern is more generally true. If we have n different measurement results for generic operators,

$$A(t)\big|a_k, t\big\rangle = \big|a_k, t\big\rangle a_k \qquad \text{for} \quad k = 1, 2, \ldots, n\,, \tag{3.7.14}$$

and consider the Schrödinger equation for the column of bras $\langle a_k, t|$,

$$\mathrm{i}\hbar\frac{\partial}{\partial t}\begin{pmatrix}\langle a_1,t|\\ \vdots\\ \langle a_n,t|\end{pmatrix}=\begin{pmatrix}\langle a_1,t|\\ \vdots\\ \langle a_n,t|\end{pmatrix}H=\mathcal{H}\begin{pmatrix}\langle a_1,t|\\ \vdots\\ \langle a_n,t|\end{pmatrix}, \tag{3.7.15}$$

we encounter the $n\times n$ matrix

$$\mathcal{H}=\begin{pmatrix}\langle a_1,t|H|a_1,t\rangle & \cdots & \langle a_1,t|H|a_n,t\rangle\\ \langle a_2,t|H|a_1,t\rangle & \cdots & \langle a_2,t|H|a_n,t\rangle\\ \vdots & \ddots & \vdots\\ \langle a_n,t|H|a_1,t\rangle & \cdots & \langle a_n,t|H|a_n,t\rangle\end{pmatrix}, \tag{3.7.16}$$

so that the probability amplitudes of an arbitrary ket

$$|\ \rangle=|a_1,t\rangle\alpha_1(t)+|a_2,t\rangle\alpha_2(t)+\cdots+|a_n,t\rangle\alpha_n(t)\mathrel{\widehat{=}}\begin{pmatrix}\alpha_1(t)\\ \alpha_2(t)\\ \vdots\\ \alpha_n(t)\end{pmatrix} \tag{3.7.17}$$

obey the Schrödinger equation

$$\mathrm{i}\hbar\frac{\partial}{\partial t}\begin{pmatrix}\alpha_1(t)\\ \alpha_2(t)\\ \vdots\\ \alpha_n(t)\end{pmatrix}=\mathcal{H}\begin{pmatrix}\alpha_1(t)\\ \alpha_2(t)\\ \vdots\\ \alpha_n(t)\end{pmatrix}. \tag{3.7.18}$$

One often writes $\psi(t)$ for such a column of probability amplitudes,

$$\psi(t)=\begin{pmatrix}\alpha_1(t)\\ \vdots\\ \alpha_n(t)\end{pmatrix}, \tag{3.7.19}$$

and then this Schrödinger equation acquires a form,

$$\mathrm{i}\hbar\frac{\partial}{\partial t}\psi=\mathcal{H}\psi\,, \tag{3.7.20}$$

that is as compact as the original statement (3.1.23) about bras.

3.8 Time-independent Schrödinger equation

If $\mathcal{H}$ has no parametric time dependence and is, therefore, constant in time, it is easy to integrate these Schrödinger equations formally, with the outcome

$$\langle \dots, t| = \langle \dots, 0| \, \mathrm{e}^{-\mathrm{i}Ht/\hbar} \tag{3.8.1}$$

for the original bra equation (3.1.23) and

$$\psi(t) = \mathrm{e}^{-\mathrm{i}\mathcal{H}t/\hbar}\psi(0) \tag{3.8.2}$$

for its numerical analog (3.7.20) for columns of probability amplitudes. Again, do not fail to note that — naturally and properly — the operator stands to the right of the bra, whereas the matrix stands to the left of the column.

Since H is hermitian, $H = H^\dagger$, the operator $\mathrm{e}^{-\mathrm{i}Ht/\hbar}$ is unitary, as it must be to conserve the unit length of the bra. But unless one can do something more explicit about this exponential function of the Hamilton operator, this formal solution of the Schrödinger equation is not of much help. Now, if we know the eigenkets and eigenbras of H,

$$H(t)|E_k, t\rangle = |E_k, t\rangle E_k \,, \qquad \langle E_k, t|H(t) = E_k\langle E_k, t| \,, \tag{3.8.3}$$

labeled by the eigenvalues E_k, the *eigenenergies* of the system, we can write

$$f(H) = \sum_k |E_k, t\rangle f(E_k) \langle E_k, t| \tag{3.8.4}$$

for any function of H, in particular that unitary exponential function. In (3.8.3), $H(t)$ is actually independent of time t, and its eigenvalues E_k are constant in time as well.

Regarding the time dependence of the eigenbras and eigenkets of H, we note that it is elementary because

$$\mathrm{i}\hbar\frac{\partial}{\partial t}\langle E_k, t| = \langle E_k, t|H = E_k\langle E_k, t| \tag{3.8.5}$$

gives

$$\langle E_k, t| = \mathrm{e}^{-\mathrm{i}E_k t/\hbar}\langle E_k, 0| \tag{3.8.6}$$

and

$$|E_k, t\rangle = |E_k, 0\rangle \, \mathrm{e}^{\mathrm{i}E_k t/\hbar} \tag{3.8.7}$$

immediately. As a consequence, the projector $|E_k,t\rangle\langle E_k,t|$ does not depend on time at all, so that we can just use it at some reference time, which we leave implicit, any instant can serve the purpose equally well. So, just writing $|E_k\rangle\langle E_k|$ for $|E_k,t\rangle\langle E_k,t|$, we have

$$f(H) = \sum_k \big|E_k\big\rangle f(E_k) \big\langle E_k\big| \tag{3.8.8}$$

with

$$H\big|E_k\big\rangle = \big|E_k\big\rangle E_k \tag{3.8.9}$$

and arrive at

$$\mathrm{e}^{-\mathrm{i}Ht/\hbar} = \sum_k \big|E_k\big\rangle\, \mathrm{e}^{-\mathrm{i}E_k t/\hbar} \big\langle E_k\big| \,. \tag{3.8.10}$$

Quite analogously, we have

$$\mathrm{e}^{-\mathrm{i}\mathcal{H}t/\hbar} = \sum_k \phi_k\, \mathrm{e}^{-\mathrm{i}E_k t/\hbar} \phi_k^\dagger \,, \tag{3.8.11}$$

where ϕ_k is the column representing ket $|E_k\rangle$ and its adjoint $\phi_k^\dagger$ is the row for bra $\langle E_k|$. They are, of course, the respective eigencolumns and eigenrows of the matrix $\mathcal{H}$,

$$\mathcal{H}\phi_k = \phi_k E_k \,, \qquad \phi_k^\dagger \mathcal{H} = E_k \phi_k^\dagger \,, \tag{3.8.12}$$

which are the numerical analogs of (3.8.3).

The eigenvalue equation

$$H\big|E_k\big\rangle = \big|E_k\big\rangle E_k \tag{3.8.13}$$

is called *time-independent Schrödinger equation*. It is a central problem to determine the eigenvalues and eigenstates of the dynamics of a system because the energy eigenvalues determine the frequencies that govern the evolution. It is clear on the outset that the E_ks themselves are of no particular physical significance because we can always change their value by adding a constant to the Hamilton operator; see Section 3.3. Only energy differences have physical meaning.

To illustrate this point once more, consider a system prepared such that its properties are described by the ket

$$\big|\ \big\rangle = \sum_k \big|E_k\big\rangle \alpha_k \,, \tag{3.8.14}$$

and we are asking for the probability of finding eigenvalue a_j of observable A at time t. This probability is the absolute square of the probability amplitude

$$\begin{aligned}\langle a_j, t| \,\rangle &= \langle a_j, 0| \,\mathrm{e}^{-\mathrm{i}Ht/\hbar} |\,\rangle \\ &= \sum_k \langle a_j, 0| \,\mathrm{e}^{-\mathrm{i}Ht/\hbar} |E_k\rangle \alpha_k \\ &= \sum_k \langle a_j, 0|E_k\rangle \,\mathrm{e}^{-\mathrm{i}E_k t/\hbar} \alpha_k \,, \end{aligned} \tag{3.8.15}$$

that is,

$$\begin{aligned}\left|\langle a_j, t| \,\rangle\right|^2 &= \sum_{k,l} \alpha_l^* \,\mathrm{e}^{\mathrm{i}E_l t/\hbar} \langle E_l|a_j, 0\rangle \langle a_j, 0|E_k\rangle \,\mathrm{e}^{-\mathrm{i}E_k t/\hbar} \alpha_k \\ &= \sum_{k,l} \alpha_l^* \langle E_l|a_j, 0\rangle \langle a_j, 0|E_k\rangle \alpha_k \,\mathrm{e}^{-\mathrm{i}(E_k - E_l)t/\hbar} \,. \end{aligned} \tag{3.8.16}$$

We read off that the (circular) frequencies of the system are

$$\omega_{kl} = \frac{E_k - E_l}{\hbar} \,, \tag{3.8.17}$$

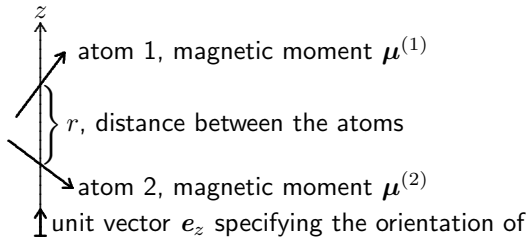
which appear in the periodic functions

$$\mathrm{e}^{-\mathrm{i}(E_k - E_l)t/\hbar} = \mathrm{e}^{-\mathrm{i}\omega_{kl}t} = \cos(\omega_{kl}t) - \mathrm{i}\sin(\omega_{kl}t) \,. \tag{3.8.18}$$

Since they refer to the difference in energy between two states of definite energy, the ω_{kl}s are often called *transition frequencies*.

3.9 Example: Two magnetic silver atoms

As an illustrating example, of quite some physical relevance for certain modern applications, consider two magnetic silver atoms, located on the z axis, a fixed distance r apart:


z
atom 1, magnetic moment μ(1)
r, distance between the atoms
atom 2, magnetic moment μ(2)
unit vector e_z specifying the orientation of
the line connecting the atomic positions

(3.9.1)

Their magnetic interaction energy is the dipole–dipole interaction energy

$$E_{\text{magn}} = \frac{\boldsymbol{\mu}^{(1)} \cdot \boldsymbol{\mu}^{(2)} - 3\boldsymbol{\mu}^{(1)} \cdot \boldsymbol{e}_z \, \boldsymbol{e}_z \cdot \boldsymbol{\mu}^{(2)}}{r^3} \,. \tag{3.9.2}$$

With $\boldsymbol{\mu}^{(1)} \propto \boldsymbol{\sigma}^{(1)}$, $\boldsymbol{\mu}^{(2)} \propto \boldsymbol{\sigma}^{(2)}$, this gives the Hamilton operator

$$H = \frac{1}{2}\hbar\omega\Big(\boldsymbol{\sigma}^{(1)} \cdot \boldsymbol{\sigma}^{(2)} - 3\sigma_z^{(1)}\sigma_z^{(2)}\Big) \,, \tag{3.9.3}$$

where the prefactor $\frac{1}{2}\hbar\omega$, of metrical dimension energy, summarizes the various constants in a convenient way. The energy unit $\hbar\omega$ introduces a corresponding frequency unit ω — or rather $\omega/(2\pi)$ because ω itself is a circular frequency, whereas $\omega/(2\pi)$ is a true frequency.

The dipole–dipole energy E_{magn} is such that the "head to tail" configurations,

$$\begin{matrix}\uparrow\\ \uparrow\end{matrix} \quad \text{and} \quad \begin{matrix}\downarrow\\ \downarrow\end{matrix} \qquad \text{both up or both down along } z, \tag{3.9.4}$$

are of minimal energy, $1 - 3 = -2$ units of energy, whereas the "head to head" or "tail to tail" configurations,

$$\begin{matrix}\downarrow\\ \uparrow\end{matrix} \quad \text{and} \quad \begin{matrix}\uparrow\\ \downarrow\end{matrix} \qquad \text{one up, the other down along } z, \tag{3.9.5}$$

are of much larger energy: $-1 + 3 = +2$ units. Accordingly, we expect $\uparrow\uparrow$ and $\downarrow\downarrow$ to be stable configurations and $\uparrow\downarrow$ or $\downarrow\uparrow$ to be unstable.

These considerations invite the use of $|\uparrow_z\uparrow_z\rangle$, $|\uparrow_z\downarrow_z\rangle$, $|\downarrow_z\uparrow_z\rangle$, $|\downarrow_z\downarrow_z\rangle$ as the reference kets to which we refer all kets, operators, *et cetera*. Thus,

$$|\,\rangle = \Big(|\uparrow_z\uparrow_z\rangle \; |\uparrow_z\downarrow_z\rangle \; |\downarrow_z\uparrow_z\rangle \; |\downarrow_z\downarrow_z\rangle\Big) \underbrace{\begin{pmatrix}\psi_1\\ \psi_2\\ \psi_3\\ \psi_4\end{pmatrix}}_{\equiv\,\psi} \mathrel{\widehat{=}} \psi \tag{3.9.6}$$

for any arbitrary ket, with the particular examples,

$$\big|\uparrow_z\uparrow_z\big\rangle \mathrel{\widehat{=}} \begin{pmatrix}1\\0\\0\\0\end{pmatrix}, \qquad \big|\uparrow_z\downarrow_z\big\rangle \mathrel{\widehat{=}} \begin{pmatrix}0\\1\\0\\0\end{pmatrix},$$
$$\big|\downarrow_z\uparrow_z\big\rangle \mathrel{\widehat{=}} \begin{pmatrix}0\\0\\1\\0\end{pmatrix}, \qquad \big|\downarrow_z\downarrow_z\big\rangle \mathrel{\widehat{=}} \begin{pmatrix}0\\0\\0\\1\end{pmatrix}, \tag{3.9.7}$$

for the chosen set of basis kets. All of them are eigenkets of $\sigma_z^{(1)}\sigma_z^{(2)}$ with eigenvalues of $(+1)(+1) = +1$, $(+1)(-1) = -1$, $(-1)(+1) = -1$, $(-1)(-1) = +1$, respectively, so that

$$\sigma_z^{(1)}\sigma_z^{(2)} \mathrel{\widehat{=}} \begin{pmatrix}1&0&0&0\\0&-1&0&0\\0&0&-1&0\\0&0&0&1\end{pmatrix} \tag{3.9.8}$$

is the resulting matrix representation for $\sigma_z^{(1)}\sigma_z^{(2)}$. Regarding $\boldsymbol{\sigma}^{(1)}\cdot\boldsymbol{\sigma}^{(2)}$, we recall what we found in Section 2.18, namely that

$$\frac{1}{2}\Big(\big|\uparrow_z\downarrow_z\big\rangle - \big|\downarrow_z\uparrow_z\big\rangle\Big)\Big(\big\langle\uparrow_z\downarrow_z\big| - \big\langle\downarrow_z\uparrow_z\big|\Big) = \frac{1}{4}\big(1 - \boldsymbol{\sigma}^{(1)}\cdot\boldsymbol{\sigma}^{(2)}\big), \tag{3.9.9}$$

so that

$$\begin{aligned}\boldsymbol{\sigma}^{(1)}\cdot\boldsymbol{\sigma}^{(2)} &= 1 - 2\Big(\big|\uparrow_z\downarrow_z\big\rangle - \big|\downarrow_z\uparrow_z\big\rangle\Big)\Big(\big\langle\uparrow_z\downarrow_z\big| - \big\langle\downarrow_z\uparrow_z\big|\Big)\\ &\mathrel{\widehat{=}} \begin{pmatrix}1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&1\end{pmatrix} - 2\begin{pmatrix}0\\1\\-1\\0\end{pmatrix}\begin{pmatrix}0&1&-1&0\end{pmatrix}\\ &= \begin{pmatrix}1&0&0&0\\0&-1&2&0\\0&2&-1&0\\0&0&0&1\end{pmatrix}\end{aligned} \tag{3.9.10}$$

is the matrix representation for $\boldsymbol{\sigma}^{(1)}\cdot\boldsymbol{\sigma}^{(2)}$.

Putting the ingredients together, we have

$$H \mathrel{\widehat{=}} \mathcal{H} = \frac{1}{2}\hbar\omega \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 2 & 0 \\ 0 & 2 & -1 & 0 \\ 0 & 0 & 0 & \end{pmatrix} - \frac{3}{2}\hbar\omega \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$
$$= \hbar\omega \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \tag{3.9.11}$$

for the matrix representing the Hamilton operator (3.9.3). The time-independent Schrödinger equation $H|E_k\rangle = |E_k\rangle E_k$ translates into the matrix eigenvalue equation

$$\mathcal{H}\psi = E\psi\,, \tag{3.9.12}$$

the solution of which can be found by simple inspection,

$$\mathcal{H}\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} = -\hbar\omega \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} : \text{ eigenvalue } -\hbar\omega, \text{ eigencolumn } \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix},$$
$$\mathcal{H}\begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} = -\hbar\omega \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} : \text{ eigenvalue } -\hbar\omega, \text{ eigencolumn } \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix},$$
$$\mathcal{H}\begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix} = 2\hbar\omega \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix} : \text{ eigenvalue } 2\hbar\omega, \text{ eigencolumn } \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix},$$
$$\mathcal{H}\begin{pmatrix} 0 \\ 1 \\ -1 \\ 0 \end{pmatrix} = 0 : \text{ eigenvalue } 0, \text{ eigencolumn } \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ 1 \\ -1 \\ 0 \end{pmatrix}. \tag{3.9.13}$$

The factors of $1/\sqrt{2}$ are for the normalization to unit length.

The general statement (3.8.8) about a function of operator H has the numerical analog

$$f(\mathcal{H}) = \sum_k \phi_k f(E_k) \phi_k^\dagger\,, \tag{3.9.14}$$

where, as in (3.8.11) and (3.8.12), ϕ_k is the column for ket $|E_k\rangle$ and its adjoint $\phi_k^\dagger$ is the row for bra $\langle E_k|$. In this example, we thus have

$$f(\mathcal{H}) = f(-\hbar\omega)\begin{pmatrix}1\\0\\0\\0\end{pmatrix}\begin{pmatrix}1&0&0&0\end{pmatrix} + f(-\hbar\omega)\begin{pmatrix}0\\0\\0\\1\end{pmatrix}\begin{pmatrix}0&0&0&1\end{pmatrix}$$
$$+ f(2\hbar\omega)\frac{1}{2}\begin{pmatrix}0\\1\\1\\0\end{pmatrix}\begin{pmatrix}0&1&1&0\end{pmatrix} + f(0)\frac{1}{2}\begin{pmatrix}0\\1\\-1\\0\end{pmatrix}\begin{pmatrix}0&1&-1&0\end{pmatrix} \quad (3.9.15)$$

or

$$f(\mathcal{H}) = \begin{pmatrix} f(-\hbar\omega) & 0 & 0 & 0 \\ 0 & \frac{1}{2}[f(0)+f(2\hbar\omega)] & \frac{1}{2}[-f(0)+f(2\hbar\omega)] & 0 \\ 0 & \frac{1}{2}[-f(0)+f(2\hbar\omega)] & \frac{1}{2}[f(0)+f(2\hbar\omega)] & 0 \\ 0 & 0 & 0 & f(-\hbar\omega) \end{pmatrix}. \quad (3.9.16)$$

We need it for $f(\mathcal{H}) = \mathrm{e}^{-\mathrm{i}\mathcal{H}t/\hbar}$, that is,

$$\begin{aligned} f(-\hbar\omega) &= \mathrm{e}^{\mathrm{i}\omega t}\,, \\ \frac{1}{2}[f(0)+f(2\hbar\omega)] &= \mathrm{e}^{-\mathrm{i}\omega t}\cos(\omega t)\,, \\ \frac{1}{2}[-f(0)+f(2\hbar\omega)] &= -\mathrm{i}\,\mathrm{e}^{-\mathrm{i}\omega t}\sin(\omega t)\,, \end{aligned} \quad (3.9.17)$$

thus

$$\mathrm{e}^{-\mathrm{i}Ht/\hbar} \mathrel{\widehat{=}} \mathrm{e}^{-\mathrm{i}\mathcal{H}t/\hbar} = \begin{pmatrix} \mathrm{e}^{\mathrm{i}\omega t} & 0 & 0 & 0 \\ 0 & \mathrm{e}^{-\mathrm{i}\omega t}\cos(\omega t) & -\mathrm{i}\,\mathrm{e}^{-\mathrm{i}\omega t}\sin(\omega t) & 0 \\ 0 & -\mathrm{i}\,\mathrm{e}^{-\mathrm{i}\omega t}\sin(\omega t) & \mathrm{e}^{-\mathrm{i}\omega t}\cos(\omega t) & 0 \\ 0 & 0 & 0 & \mathrm{e}^{\mathrm{i}\omega t} \end{pmatrix}. \quad (3.9.18)$$

In particular, then, we get

$$\psi(t=0) = \begin{pmatrix}1\\0\\0\\0\end{pmatrix} \to \psi(t) = \mathrm{e}^{-\mathrm{i}\mathcal{H}t/\hbar}\begin{pmatrix}1\\0\\0\\0\end{pmatrix} = \begin{pmatrix}\mathrm{e}^{\mathrm{i}\omega t}\\0\\0\\0\end{pmatrix} = \mathrm{e}^{\mathrm{i}\omega t}\psi(0), \quad (3.9.19)$$

which, translated back into a statement about kets, reads

$$\underbrace{\left|\uparrow_z\uparrow_z, t=0\right\rangle}_{\substack{\text{time-independent}\\ \text{state ket}}} = \underbrace{\left|\uparrow_z\uparrow_z, t\right\rangle}_{\substack{\text{time-dependent}\\ \text{basis ket}}} \underbrace{\mathrm{e}^{\mathrm{i}\omega t}}_{\substack{\text{time-dependent}\\ \text{probability}\\ \text{amplitude}}}, \tag{3.9.20}$$

where we emphasize once more that the ket on the left, which specifies the state of the two-atom system, refers to the initial time $t = 0$ and, therefore, has no actual time dependence and that the time dependence of the ket and probability amplitude on the right compensate for each other.

Thus, atoms that are initially aligned, $\uparrow_z\uparrow_z$, remain aligned as time passes. The same is true for $\downarrow_z\downarrow_z$, the other stable configuration, inasmuch as

$$\psi(0) = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \rightarrow \psi(t) = \mathrm{e}^{-\mathrm{i}\mathcal{H}t/\hbar}\psi(0) = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \mathrm{e}^{\mathrm{i}\omega t} \end{pmatrix} = \mathrm{e}^{\mathrm{i}\omega t}\psi(0)\,. \tag{3.9.21}$$

But the unstable configurations, $\uparrow_z\downarrow_z$ and $\downarrow_z\uparrow_z$, do evolve in time into other arrangements. For example, if we have $\uparrow_z\downarrow_z$ initially, that is,

$$\psi(0) = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \tag{3.9.22}$$

we get

$$\psi(t) = \mathrm{e}^{-\mathrm{i}\mathcal{H}t/\hbar}\psi(0) = \mathrm{e}^{-\mathrm{i}\omega t} \begin{pmatrix} 0 \\ \cos(\omega t) \\ -\mathrm{i}\sin(\omega t) \\ 0 \end{pmatrix}, \tag{3.9.23}$$

stating that the probability of $\uparrow_z\downarrow_z$ at time t is

$$\left|\mathrm{e}^{-\mathrm{i}\omega t}\cos(\omega t)\right|^2 = \cos(\omega t)^2 \tag{3.9.24}$$

and that of $\downarrow_z\uparrow_z$ at time t is

$$\left|\mathrm{e}^{-\mathrm{i}\omega t}(-\mathrm{i})\sin(\omega t)\right|^2 = \sin(\omega t)^2\,, \tag{3.9.25}$$

whereas the probabilities of $\uparrow_z\uparrow_z$ and $\downarrow_z\downarrow_z$ are zero for all times. So, if we wait for time $t = \pi/(2\omega)$, then $\uparrow_z\downarrow_z$ has been completely turned into $\downarrow_z\uparrow_z$, and waiting for another lapse of $\pi/(2\omega)$, thus reaching $t = \pi/\omega$, turns it back into $\uparrow_z\downarrow_z$. There is a periodic transition from $\uparrow_z\downarrow_z$ to $\downarrow_z\uparrow_z$ and back, with superpositions at intermediate times. For instance, at time $t = \pi/(4\omega)$, we have

$$\psi\Big(t = \frac{\pi}{4\omega}\Big) = \mathrm{e}^{-\mathrm{i}\pi/4} \begin{pmatrix} 0 \\ \cos(\pi/4) \\ -\mathrm{i}\sin(\pi/4) \\ 0 \end{pmatrix} = \frac{1-\mathrm{i}}{2} \begin{pmatrix} 0 \\ 1 \\ -\mathrm{i} \\ 0 \end{pmatrix}, \tag{3.9.26}$$

which translates into

$$\big|\uparrow_z\downarrow_z, 0\big\rangle = \Big(\big|\uparrow_z\downarrow_z, \frac{\pi}{4\omega}\big\rangle - \mathrm{i}\big|\downarrow_z\uparrow_z, \frac{\pi}{4\omega}\big\rangle\Big)\frac{1-\mathrm{i}}{2}\,. \tag{3.9.27}$$

The frequency at which these oscillations happen is revealed upon writing the probabilities as

$$\left.\begin{matrix}\cos(\omega t)^2 \\ \sin(\omega t)^2\end{matrix}\right\} = \frac{1}{2} \pm \frac{1}{2}\cos(2\omega t) \tag{3.9.28}$$

so that $2\omega = (2\hbar\omega - 0\hbar\omega)/\hbar$ is the transition frequency, correctly related to the energy difference between the eigenstates of H that are involved, see

$$\big|\uparrow_z\downarrow_z\big\rangle = \frac{1}{\sqrt{2}}\Bigg(\underbrace{\frac{1}{\sqrt{2}}\Big(\big|\uparrow_z\downarrow_z\big\rangle + \big|\downarrow_z\uparrow_z\big\rangle\Big)}_{\text{energy eigenvalue } 2\hbar\omega} + \underbrace{\frac{1}{\sqrt{2}}\Big(\big|\uparrow_z\downarrow_z\big\rangle - \big|\downarrow_z\uparrow_z\big\rangle\Big)}_{\text{energy eigenvalue } 0\hbar\omega}\Bigg)\,. \tag{3.9.29}$$

Having the machinery set up, we can also ask more complicated questions, such as the following: If we have $\uparrow_x\uparrow_x$ at the initial time $t = 0$, what is the probability of $\uparrow_x\uparrow_x$ at the later time t? We begin with noting that

$$\big|\uparrow_x\uparrow_x\big\rangle = \frac{1}{2}\Big(\big|\uparrow_z\uparrow_z\big\rangle + \big|\uparrow_z\downarrow_z\big\rangle + \big|\downarrow_z\uparrow_z\big\rangle + \big|\downarrow_z\downarrow_z\big\rangle\Big) \tag{3.9.30}$$

so that

$$\psi(0) = \frac{1}{2}\begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} \tag{3.9.31}$$

and

$$\psi(t) = \mathrm{e}^{-\mathrm{i}\mathcal{H}t/\hbar}\psi(0) = \frac{1}{2}\begin{pmatrix} \mathrm{e}^{\mathrm{i}\omega t} \\ \mathrm{e}^{-2\mathrm{i}\omega t} \\ \mathrm{e}^{-2\mathrm{i}\omega t} \\ \mathrm{e}^{\mathrm{i}\omega t} \end{pmatrix}. \tag{3.9.32}$$

The probability amplitude $\langle\uparrow_x\uparrow_x, t|\uparrow_x\uparrow_x, 0\rangle$ is then

$$\underbrace{\frac{1}{2}\begin{pmatrix}1&1&1&1\end{pmatrix}}_{\langle\uparrow_x\uparrow_x,t|\,\widehat{=}}\underbrace{\psi(t)}_{\widehat{=}\,|\uparrow_x\uparrow_x,0\rangle} = \frac{1}{2}\Big(\mathrm{e}^{\mathrm{i}\omega t} + \mathrm{e}^{-2\mathrm{i}\omega t}\Big) = \mathrm{e}^{-\frac{1}{2}\mathrm{i}\omega t}\cos\Big(\frac{3}{2}\omega t\Big) \tag{3.9.33}$$

so that the transition probability

$$\begin{aligned}\mathrm{prob}(\uparrow_x\uparrow_x \text{ at } t=0 \longrightarrow \uparrow_x\uparrow_x \text{ at } t) &= \cos\Big(\frac{3}{2}\omega t\Big)^2 \\ &= \frac{1}{2} + \frac{1}{2}\cos(3\omega t)\end{aligned} \tag{3.9.34}$$

oscillates with frequency 3ω. This is as it should be because the decomposition

$$|\uparrow_x\uparrow_x\rangle = \frac{1}{\sqrt{2}}\Bigg(\underbrace{\frac{1}{\sqrt{2}}\Big(|\uparrow_z\uparrow_z\rangle + |\downarrow_z\downarrow_z\rangle\Big)}_{\text{energy} = -\hbar\omega} + \underbrace{\frac{1}{\sqrt{2}}\Big(|\uparrow_z\downarrow_z\rangle + |\downarrow_z\uparrow_z\rangle\Big)}_{\text{energy} = 2\hbar\omega}\Bigg) \tag{3.9.35}$$

shows that the eigenstates of H with the energies $-\hbar\omega$ and $2\hbar\omega$ are involved so that

$$\frac{1}{\hbar}\big(2\hbar\omega - (-\hbar\omega)\big) = 3\omega \tag{3.9.36}$$

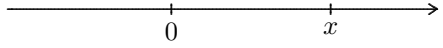
is the frequency associated with the energy difference.

Chapter 4

Motion Along the x Axis

4.1 Kets, bras, and wave functions

At the beginning of Section 2.21, we noted the possibility of measurements that have a continuum of possible measurement results, rather than a finite number of outcomes. The prime example is the position of an atom along a line — the x axis, say,



(4.1.1)

We look for the atom and find it at x but not at x'. Following the structure established for discrete observables A and their eigenkets $|a_k\rangle$ to measurement result a_k, we now introduce ket $|x\rangle$ for "atom at position x," and its bra $\langle x| = |x\rangle^\dagger$, and note

$$\langle x|x'\rangle = 0 \quad \text{if} \quad x \neq x' \,, \tag{4.1.2}$$

saying the following: "an atom at x' is not at x" but do not specify right now the right-hand side for $x = x'$.

The analog of

$$|\ \rangle = \sum_k |a_k\rangle\langle a_k|\ \rangle = \sum_k |a_k\rangle\psi_k \tag{4.1.3}$$

with the *probability amplitudes* $\psi_k = \langle a_k|\ \rangle$ is then

$$|\ \rangle = \int \mathrm{d}x\, |x\rangle\langle x|\ \rangle = \int \mathrm{d}x\, |x\rangle\psi(x) \tag{4.1.4}$$

with the *wave function* $\psi(x) = \langle x|\ \rangle$, where, fitting to the continuous character of x, we perform an integration as the continuum analog of a summation.

Then, as the analogs of

$$\langle 1|2\rangle = \sum_k \langle 1|a_k\rangle\langle a_k|2\rangle = \sum_k \psi_k^{(1)*}\psi_k^{(2)}\,, \tag{4.1.5}$$

we have

$$\langle 1|2\rangle = \int \mathrm{d}x\,\langle 1|x\rangle\langle x|2\rangle = \int \mathrm{d}x\,\psi^{(1)}(x)^*\psi^{(2)}(x) \tag{4.1.6}$$

or

$$\langle 1|2\rangle = \langle 1|\left(\int \mathrm{d}x\,|x\rangle\langle x|\right)|2\rangle\,. \tag{4.1.7}$$

Accordingly, the *completeness* of the $|x\rangle$ kets and the $\langle x|$ bras is expressed by

$$\int \mathrm{d}x\,|x\rangle\langle x| = 1\,, \tag{4.1.8}$$

the analog of

$$\sum_k |a_k\rangle\langle a_k| = 1\,. \tag{4.1.9}$$

The square of the identity is the identity itself, $1^2 = 1$. In the discrete case, this translates into

$$\sum_j |a_j\rangle\langle a_j| \sum_k |a_k\rangle\langle a_k| = \sum_j |a_j\rangle\langle a_j|\,, \tag{4.1.10}$$

which is a familiar consequence of the orthonormality of the kets $|a_k\rangle$,

$$\langle a_j|a_k\rangle = \delta_{jk}\,. \tag{4.1.11}$$

For the continuum of $|x\rangle$ kets, we need the appropriate generalization of δ_{jk}, the discrete Kronecker delta symbol, the right-hand side of $\langle x|x'\rangle = \cdots$ to ensure that $1^2 = 1$ translates into

$$\begin{aligned}\int \mathrm{d}x\,|x\rangle\langle x| \int \mathrm{d}x'\,|x'\rangle\langle x'| &= \int \mathrm{d}x\,|x\rangle \underbrace{\int \mathrm{d}x'\,\langle x|x'\rangle\langle x'|}_{=\langle x|\ \text{(needed)}} \\ &= \int \mathrm{d}x\,|x\rangle\langle x|\,. \end{aligned} \tag{4.1.12}$$

Following Dirac, we write

$$\langle x|x'\rangle = \delta(x - x')\,, \tag{4.1.13}$$

which states the orthonormality of the continuum of position states in close analogy to the discrete statement (4.1.11). On the right-hand side, we have the so-called *Dirac delta function*, whose defining property is

$$\int \mathrm{d}x'\, \delta(x - x') f(x') = f(x) \tag{4.1.14}$$

for all continuous functions $f(x)$. Clearly, this is the continuum analog of

$$\sum_k \delta_{jk} f_k = f_j \tag{4.1.15}$$

in all respects, except for one: whereas the values of δ_{jk} form an ordinary row for each j, and an ordinary column for each k, the Dirac delta function is no ordinary function of x' for any fixed value of x because it would have to possess extraordinary, yes contradictory, properties: We would need $\delta(x - x') = 0$ if $x' \neq x$ and at the same time

$$\int \mathrm{d}x'\, \delta(x - x') = 1\,, \tag{4.1.16}$$

that is, unit area below a curve that coincides with the x' axis everywhere except at $x' = x$. No ordinary function can be like this.

In fact, Dirac's delta function is not a function in the usual sense — mathematically speaking, it is a distribution — but rather an integral kernel that implements the mapping

$$\begin{aligned} \text{function } f &\longrightarrow \text{ its value } f(x) \text{ at } x, \\ f &\longrightarrow \int \mathrm{d}x'\, \delta(x - x') f(x') = f(x)\,. \end{aligned} \tag{4.1.17}$$

For this to be a well-defined procedure, the function $x' \mapsto f(x')$ must have a well-defined value at $x' = x$, which is to say that it should be continuous there. Otherwise, it could matter whether we approach x from the left or the right.

It is often expedient to regard the Dirac delta function as the limit of ordinary functions that are very strongly peaked at the distinguished point.

The general recipe is as follows:

(a) Take a function $D(\phi)$ with the properties

$$\begin{aligned} &\text{(i)}\ \ D(\phi) \to 0 \quad \text{as } |\phi| \to \infty\,, \\ &\text{(ii)}\ D(\phi) \to 1 \quad \text{as } \phi \to 0\,, \end{aligned}$$

(b) define, for $\epsilon > 0$,

$$\delta(x - x'; \epsilon) = \int \frac{\mathrm{d}k}{2\pi}\, D(\epsilon k)\, \mathrm{e}^{\mathrm{i}k(x\,-\,x')}\,,$$

(c) then

$$\delta(x - x') = \lim_{\epsilon \to 0} \delta(x - x'; \epsilon) \tag{4.1.18}$$

in the sense of

$$f(x) = \lim_{\epsilon \to 0} \int \mathrm{d}x'\, \delta(x - x'; \epsilon) f(x')\,. \tag{4.1.19}$$

We say that such an ordinary function $\delta(x - x'; \epsilon)$ is a *model* for the Dirac delta function.

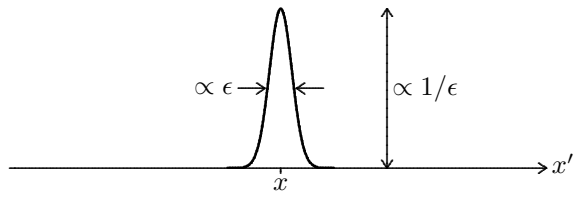
A simple example is

$$D(\phi) = \mathrm{e}^{-|\phi|}\,, \tag{4.1.20}$$

for which

$$\begin{aligned} \delta(x - x'; \epsilon) &= \int \frac{\mathrm{d}k}{2\pi}\, \mathrm{e}^{-|\epsilon k|}\, \mathrm{e}^{\mathrm{i}k(x\,-\,x')} \\ &= \frac{1}{\pi}\, \mathrm{Re}\left(\int_0^\infty \mathrm{d}k\ \mathrm{e}^{-\epsilon k\, +\, \mathrm{i}k(x\,-\,x')} \right) \\ &= \frac{1}{\pi}\, \mathrm{Re}\left(\frac{1}{\epsilon - \mathrm{i}(x - x')} \right) \\ &= \frac{1}{\pi}\, \frac{\epsilon}{\epsilon^2 + (x - x')^2}\,. \end{aligned} \tag{4.1.21}$$

Its rough graphical representation is typical:

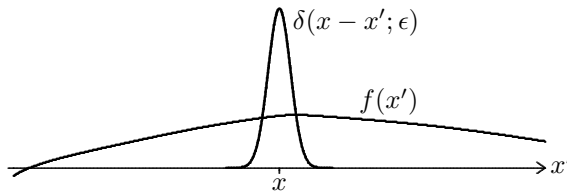


(4.1.22)

a very narrow, very high peak centered at $x = x'$. The area under the function is

$$\begin{aligned}\int \mathrm{d}x'\, \delta(x - x'; \epsilon) &= \frac{1}{\pi} \int \mathrm{d}x'\, \frac{\epsilon}{\epsilon^2 + (x - x')^2} \\ &= \frac{1}{\pi} \tan^{-1}\Big(\frac{x' - x}{\epsilon}\Big)\Big|_{x' = -\infty}^{\infty} \\ &= \frac{1}{\pi}\left(\frac{\pi}{2} - \left(-\frac{\pi}{2}\right)\right) = 1\,, \end{aligned} \tag{4.1.23}$$

as it should be. So, any continuous function $f(x')$ that changes little over the range of x' values where $\delta(x - x'; \epsilon)$ is markedly different from 0,



(4.1.24)

can be replaced by its value at x without affecting the value of the integral,

$$\int \mathrm{d}x'\, \delta(x - x'; \epsilon)\, \underbrace{f(x')}_{\cong f(x)} \cong f(x) \underbrace{\int \mathrm{d}x'\, \delta(x - x'; \epsilon)}_{=1}\,, \tag{4.1.25}$$

and thus we have a model for the Dirac delta function, indeed. Exercise 74 deals with another important model; yet another example is the subject matter of Exercise 41 in *Perturbed Evolution*.

The observation that the orthonormality of the $|x\rangle$ kets is expressed by

$$\langle x | x' \rangle = \delta(x - x') = \begin{cases} 0 & \text{if } x \neq x' \\ \infty & \text{if } x = x' \end{cases} \qquad \text{(in some sense)} \tag{4.1.26}$$

reminds us of the overidealization inherent in identifying $|x\rangle$ with "atom at position x." It overidealizes the realistic physical situation in which an absolutely precise position can never be known, because such absolute precision is an illusion, resources are always limited. You can locate an atom with any desired finite precision (given large finite resources) but not with infinite precision. In short, the kets $|x\rangle$ do not refer to any real physical state of the atom, but nevertheless, they are extremely useful mathematical entities, without which we would not be able to go about continuous degrees of freedom efficiently.

4.2 Position operator

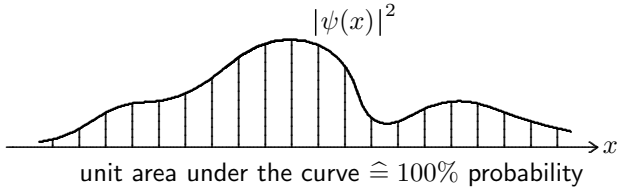
The bra-ket product of (4.1.6),

$$\langle 1|2\rangle = \int \mathrm{d}x\, \psi^{(1)}(x)^* \psi^{(2)}(x)\,, \tag{4.2.1}$$

has the normalization statement

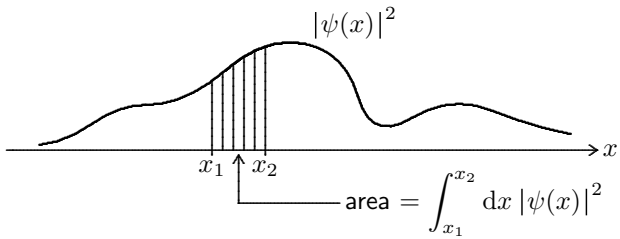
$$1 = \langle\ |\ \rangle = \int \mathrm{d}x\, \psi(x)^* \psi(x) = \int \mathrm{d}x\, |\psi(x)|^2 \tag{4.2.2}$$

as a particular case. It says that the area under the curve $|\psi(x)|^2$ is unity:



(4.2.3)

just stating that there is unit probability of finding the atom *somewhere* when we look for it. Clearly, then, any slice of the area,



(4.2.4)

is the probability of finding the atom in this x range,

$$\mathrm{prob}(x_1 < x < x_2) = \int_{x_1}^{x_2} \mathrm{d}x\, |\psi(x)|^2\,. \tag{4.2.5}$$

And so we note that $|\psi(x)|^2$ is a *probability density*, not a probability itself. It turns into a probability upon multiplication with $\mathrm{d}x$,

$$\mathrm{prob}(\text{atom in } x \cdots x + \mathrm{d}x) = \mathrm{d}x\, |\psi(x)|^2 = \langle\ |x\rangle\, \mathrm{d}x\, \langle x|\ \rangle\,. \tag{4.2.6}$$

This is the expectation value of $|x\rangle\, \mathrm{d}x\, \langle x|$, which therefore has the meaning of "projector on the infinitesimal vicinity of x." It is the infinitesimal

version of

$$\int_{x_1}^{x_2} \mathrm{d}x\, |x\rangle\langle x| = \text{projector on the range } x_1 < x < x_2\,, \tag{4.2.7}$$

whose expectation value is $\mathrm{prob}(x_1 < x < x_2)$ of (4.2.5), depicted in (4.2.4).

With these probabilities at hand, we can evaluate weighted averages, *mean values*, of x, x^2, ..., such as

$$\begin{aligned} \overline{x} = \int \mathrm{d}x\, x|\psi(x)|^2 &= \int \mathrm{d}x\, \langle\ |x\rangle x\langle x|\ \rangle \\ &= \langle\ |\left(\int \mathrm{d}x\, |x\rangle x\langle x|\right)|\ \rangle \end{aligned} \tag{4.2.8}$$

and

$$\begin{aligned} \overline{x^2} = \int \mathrm{d}x\, x^2|\psi(x)|^2 &= \int \mathrm{d}x\, \langle\ |x\rangle x^2\langle x|\ \rangle \\ &= \langle\ |\left(\int \mathrm{d}x\, |x\rangle x^2\langle x|\right)|\ \rangle \end{aligned} \tag{4.2.9}$$

and so forth. We thus identify the *position operator*

$$X = \int \mathrm{d}x\, |x\rangle x\langle x|\,, \tag{4.2.10}$$

whose expectation value is the mean value of x,

$$\langle X\rangle = \int \mathrm{d}x\, x|\psi(x)|^2\,. \tag{4.2.11}$$

This is a consistent identification because the expectation value $\langle X^2\rangle$ is really the mean value of x^2, inasmuch as

$$\begin{aligned} X^2 &= \left(\int \mathrm{d}x\, |x\rangle x\langle x|\right)\left(\int \mathrm{d}x'\, |x'\rangle x'\langle x'|\right) \\ &= \int \mathrm{d}x\, |x\rangle x \underbrace{\int \mathrm{d}x'\, \delta(x-x')x'\langle x'|}_{=x\langle x|} \\ &= \int \mathrm{d}x\, |x\rangle x^2\langle x|\,, \end{aligned} \tag{4.2.12}$$

as it should be. These generalize immediately to arbitrary functions of X,

$$f(X) = \int \mathrm{d}x\, |x\rangle f(x)\langle x|\,, \tag{4.2.13}$$

which we recognize as the continuous analog of the discrete construction

$$f(A) = \sum_k \big|a_k\big\rangle f(a_k) \big\langle a_k\big| \tag{4.2.14}$$

of (2.21.17).

The eigenvalue equation

$$f(A)\big|a_k\big\rangle = \big|a_k\big\rangle f(a_k) \tag{4.2.15}$$

has its obvious analog as well,

$$f(X)\big|x\big\rangle = \big|x\big\rangle f(x)\,. \tag{4.2.16}$$

Let us check this:

$$f(X)\big|x\big\rangle = \int \mathrm{d}x' \,\big|x'\big\rangle f(x') \underbrace{\big\langle x'\big|x\big\rangle}_{=\delta(x'-x)} = \big|x\big\rangle f(x)\,, \tag{4.2.17}$$

indeed. Everything fits together very neatly.

4.3 Momentum operator

Since the eigenkets of position operator X are labeled by the continuous parameter x, it is natural to ask what happens if we differentiate with respect to x. It is slightly more systematic to look at bras $\langle x|$ first, so we ask what is the right-hand side in

$$\frac{\partial}{\partial x}\big\langle x\big| = \textbf{?} \tag{4.3.1}$$

The outcome is a linear combination of the bras $\langle x|$, of course, and that is also the case for higher derivatives. We can talk about all of them at once by making use of Taylor's* theorem,

$$f(x+x') = \sum_{k=0}^{\infty} \frac{1}{k!}\, x'^k \left(\frac{\partial}{\partial x}\right)^k f(x)\,, \tag{4.3.2}$$

which we write compactly as

$$f(x+x') = \mathrm{e}^{x'\frac{\partial}{\partial x}} f(x)\,. \tag{4.3.3}$$

*Brook TAYLOR (1685–1731)

When applied to $\langle x|$, it reads

$$\big\langle x + x' \big| = \mathrm{e}^{x' \frac{\partial}{\partial x}} \big\langle x \big| \,. \tag{4.3.4}$$

But the mapping $\langle x| \to \langle x + x'|$ is unitary,

$$\big\langle x + x' \big| = \big\langle x \big| U \tag{4.3.5}$$

with

$$U = \int \mathrm{d}x \, \big| x \big\rangle \big\langle x \big| U = \int \mathrm{d}x \, \big| x \big\rangle \big\langle x + x' \big| \,. \tag{4.3.6}$$

See

$$\begin{aligned} U^\dagger U &= \int \mathrm{d}x'' \, \big| x'' + x' \big\rangle \underbrace{\big\langle x'' \big| \int \mathrm{d}x \, \big| x \big\rangle}_{\to \int \mathrm{d}x \, \delta(x'' - x)} \big\langle x + x' \big| \\ &= \int \mathrm{d}x'' \, \big| x'' + x' \big\rangle \big\langle x'' + x' \big| \\ &= \int \mathrm{d}x \, \big| x \big\rangle \big\langle x \big| = 1 \end{aligned} \tag{4.3.7}$$

and

$$\begin{aligned} U U^\dagger &= \int \mathrm{d}x \, \big| x \big\rangle \underbrace{\big\langle x + x' \big| \int \mathrm{d}x'' \, \big| x'' + x' \big\rangle}_{\to \int \mathrm{d}x'' \, \delta(x - x'')} \big\langle x'' \big| \\ &= \int \mathrm{d}x \, \big| x \big\rangle \big\langle x \big| = 1 \,. \end{aligned} \tag{4.3.8}$$

As a unitary operator, U is the exponential of some hermitian operator (times i), see (2.23.3), and the exponential will be linear in x', so

$$U = \mathrm{e}^{\mathrm{i} x' P/\hbar} \tag{4.3.9}$$

with $P = P^\dagger$ of metrical dimension action/distance = momentum. Indeed, this P is the *momentum operator*, an identification that will be justified in Section 4.10. Its action on position eigenstates is stated in (4.3.5), namely

$$\big\langle x \big| \, \mathrm{e}^{\mathrm{i} x' P/\hbar} = \big\langle x + x' \big| = \mathrm{e}^{x' \frac{\partial}{\partial x}} \big\langle x \big| \,. \tag{4.3.10}$$

A term-by-term comparison of an expansion in powers of x' tells us that

$$\begin{aligned}\langle x|P &= \frac{\hbar}{\mathrm{i}}\frac{\partial}{\partial x}\langle x|\,,\\ \langle x|P^2 &= \left(\frac{\hbar}{\mathrm{i}}\frac{\partial}{\partial x}\right)^2\langle x|\,,\\ \langle x|P^n &= \left(\frac{\hbar}{\mathrm{i}}\frac{\partial}{\partial x}\right)^n\langle x|\,,\end{aligned} \tag{4.3.11}$$

for $n = 0, 1, 2, \ldots$. The fundamental statement is the one for $n = 1$,

$$P \mathrel{\widehat{=}} \frac{\hbar}{\mathrm{i}}\frac{\partial}{\partial x} \qquad \text{(when acting on position bras)}, \tag{4.3.12}$$

the historical starting point of Schrödinger's formulation of quantum mechanics in terms of differential equations obeyed by the position wave function $\psi(x)$.

Note, however, the important detail that this identification of the momentum operator with the differential operator $\frac{\hbar}{\mathrm{i}}\frac{\partial}{\partial x}$ is only correct when we consider the action of P on bra $\langle x|$. In the case of P acting on ket $|x\rangle$, we have the adjoint statement

$$P|x\rangle = \mathrm{i}\hbar\frac{\partial}{\partial x}|x\rangle \tag{4.3.13}$$

and the identification would read

$$P \mathrel{\widehat{=}} \mathrm{i}\hbar\frac{\partial}{\partial x} \qquad \text{(when acting on position kets)} \tag{4.3.14}$$

and, clearly, these are just two of many identifications of this sort.

4.4 Heisenberg's commutation relation

Now consider the products XP and PX. We have

$$\langle x|XP = x\langle x|P = x\frac{\hbar}{\mathrm{i}}\frac{\partial}{\partial x}\langle x| \tag{4.4.1}$$

and

$$\langle x|PX = \frac{\hbar}{\mathrm{i}}\frac{\partial}{\partial x}\langle x|X = \frac{\hbar}{\mathrm{i}}\frac{\partial}{\partial x}\Big(x\langle x|\Big) = x\frac{\hbar}{\mathrm{i}}\frac{\partial}{\partial x}\langle x| + \frac{\hbar}{\mathrm{i}}\langle x| \tag{4.4.2}$$

so that

$$\langle x|(XP - PX) = \mathrm{i}\hbar\langle x| \tag{4.4.3}$$

for all position bras $\langle x|$. But they are complete so that

$$XP - PX = \mathrm{i}\hbar \tag{4.4.4}$$

or

$$[X, P] = \mathrm{i}\hbar\,. \tag{4.4.5}$$

This is the *Heisenberg commutator*, the famous position–momentum commutation relation, discovered by Heisenberg with quite some help from Born,* the starting point of Heisenberg's formulation of quantum mechanics in terms of operators and their evolution.

4.5 Position–momentum transformation function

The Heisenberg commutator is invariant under the exchange of X and P, in the sense of

$$X \to P, \qquad P \to -X \tag{4.5.1}$$

since

$$[X, P] \to [P, -X] = [X, P]\,, \tag{4.5.2}$$

which tells us that the properties of operator P are largely the same as those of operator X; their roles are interchangeable. As an immediate consequence, there must also be a continuum of eigenbras and eigenkets of P,

$$\langle p|P = p\langle p|\,, \quad P|p\rangle = |p\rangle p\,, \tag{4.5.3}$$

that are orthonormal (in the Dirac-delta-function sense),

$$\langle p|p'\rangle = \delta(p - p')\,, \tag{4.5.4}$$

and complete,

$$\int \mathrm{d}p\, |p\rangle\langle p| = 1\,. \tag{4.5.5}$$

To relate them to the position eigenstates

$$|p\rangle = \int \mathrm{d}x\, |x\rangle\langle x|p\rangle\,, \tag{4.5.6}$$

*Max Born (1882–1970)

we need the *transformation function* $\langle x|p\rangle$. We find it by exploiting the eigenket property,

$$\langle x|P|p\rangle = \langle x|p\rangle p\,, \tag{4.5.7}$$

and Schrödinger's differential operator identification,

$$\langle x|P|p\rangle = \frac{\hbar}{\mathrm{i}}\frac{\partial}{\partial x}\langle x|p\rangle\,, \tag{4.5.8}$$

for first establishing the differential equation

$$\frac{\partial}{\partial x}\langle x|p\rangle = \frac{\mathrm{i}}{\hbar}p\langle x|p\rangle\,, \tag{4.5.9}$$

and then solving it,

$$\langle x|p\rangle = C(p)\,\mathrm{e}^{\mathrm{i}xp/\hbar}\,, \tag{4.5.10}$$

where $C(p)$ is, for now, an arbitrary function of p.

We learn something about it upon combining the orthonormality relation of the x states with the completeness relation for the p states,

$$\begin{aligned}\delta(x-x') = \langle x|x'\rangle &= \int \mathrm{d}p\,\langle x|p\rangle\langle p|x'\rangle \\ &= \int \mathrm{d}p\,\big|C(p)\big|^2\,\mathrm{e}^{\mathrm{i}p(x-x')/\hbar} \\ &= \int \frac{\mathrm{d}k}{2\pi}\,2\pi\hbar\,\big|C(p)\big|^2\,\mathrm{e}^{\mathrm{i}k(x-x')}\,,\end{aligned} \tag{4.5.11}$$

where $k = p/\hbar$. Compare this with

$$\begin{aligned}\delta(x-x') &= \lim_{\epsilon\to 0}\delta(x-x';\epsilon) \\ &= \lim_{\epsilon\to 0}\int \frac{\mathrm{d}k}{2\pi}D(\epsilon k)\,\mathrm{e}^{\mathrm{i}k(x-x')}\end{aligned} \tag{4.5.12}$$

in (4.1.18) and conclude that consistency requires

$$2\pi\hbar\,\big|C(p)\big|^2 = \lim_{\epsilon\to 0}D(\epsilon k) = 1 \tag{4.5.13}$$

or

$$\big|C(p)\big| = \frac{1}{\sqrt{2\pi\hbar}}\,. \tag{4.5.14}$$

It is up to us to assign arbitrary p dependent phases to $|\,\rangle = |p\rangle$, but there is no benefit in deviating from the standard choice of $C(p) = \big|C(p)\big|$. Then,

$$\langle x|p\rangle = \frac{1}{\sqrt{2\pi\hbar}}\,\mathrm{e}^{\mathrm{i}xp/\hbar}\,. \tag{4.5.15}$$

Reading this as the wave function $\psi(x) = \langle x|\,\rangle$ to the ket $|\rangle = |p\rangle$, we note that a momentum eigenstate is a (plane) wave with wavelength

$$\lambda = \frac{2\pi\hbar}{p} \tag{4.5.16}$$

since an increment of x by λ changes $xp/\hbar$ by 2π and thus corresponds to one full period of this wave.

The relation $\lambda = 2\pi\hbar/p$ is known as the de Broglie* relation, and we call this λ the *de Broglie wavelength* for momentum p. It was de Broglie's work that triggered Schrödinger's interest because once you have such waves $\psi(x) \propto \mathrm{e}^{\mathrm{i}2\pi x/\lambda}$, you start wondering about the underlying wave equation — and that is what became known as the Schrödinger equation; more about it shortly.

4.6 Expectation values

By essentially reversing the above argument, we can establish the action of position operator X on momentum kets $|p\rangle$,

$$\begin{aligned}\langle x|X|p\rangle &= x\langle x|p\rangle = x\frac{1}{\sqrt{2\pi\hbar}}\,\mathrm{e}^{\mathrm{i}xp/\hbar}\\ &= \frac{\hbar}{\mathrm{i}}\frac{\partial}{\partial p}\langle x|p\rangle = \langle x|\frac{\hbar}{\mathrm{i}}\frac{\partial}{\partial p}|p\rangle\end{aligned} \tag{4.6.1}$$

so that

$$X|p\rangle = \frac{\hbar}{\mathrm{i}}\frac{\partial}{\partial p}|p\rangle \tag{4.6.2}$$

follows upon invoking the completeness of the x states once more.

Just as we did for operator X, we can speak about functions of operator P in the usual way, proceeding from

$$P = \int \mathrm{d}p\,|p\rangle p\langle p| \tag{4.6.3}$$

*Louis-Victor de Broglie (1892–1987)

and continuing with

$$\begin{aligned} P^2 &= \int \mathrm{d}p \, |p\rangle p^2 \langle p| \,, \\ P^3 &= \int \mathrm{d}p \, |p\rangle p^3 \langle p| \,, \\ &\vdots \\ P^n &= \int \mathrm{d}p \, |p\rangle p^n \langle p| \,, \end{aligned} \tag{4.6.4}$$

and finally, most generally,

$$f(P) = \int \mathrm{d}p \, |p\rangle f(p) \langle p| \,, \tag{4.6.5}$$

the obvious analog of the position version of (4.2.13),

$$f(X) = \int \mathrm{d}x \, |x\rangle f(x) \langle x| \,, \tag{4.6.6}$$

and the earlier discrete version of (2.21.17),

$$f(A) = \sum_{k=1}^{n} |a_k\rangle f(a_k) \langle a_k| \,. \tag{4.6.7}$$

Accordingly, we evaluate expectation values, or average values, of such functions by integration,

$$\begin{aligned} \big\langle f(X) \big\rangle &= \int \mathrm{d}x \, |\psi(x)|^2 f(x) \,, \\ \big\langle g(P) \big\rangle &= \int \mathrm{d}p \, |\psi(p)|^2 g(p) \,, \end{aligned} \tag{4.6.8}$$

where

$$\psi(x) = \langle x| \, \rangle \tag{4.6.9}$$

is the position wave function and

$$\psi(p) = \langle p| \, \rangle \tag{4.6.10}$$

is the momentum wave function. The transformation functions

$$\begin{aligned} \langle x|p\rangle &= \frac{1}{\sqrt{2\pi\hbar}} \, \mathrm{e}^{\mathrm{i}xp/\hbar} \,, \\ \langle p|x\rangle &= \frac{1}{\sqrt{2\pi\hbar}} \, \mathrm{e}^{-\mathrm{i}px/\hbar} \end{aligned} \tag{4.6.11}$$

relate them to each other,

$$\psi(x) = \int \mathrm{d}p \, \langle x|p\rangle\langle p| \, \rangle = \int \mathrm{d}p \, \frac{\mathrm{e}^{\mathrm{i}xp/\hbar}}{\sqrt{2\pi\hbar}}\psi(p)\,,$$
$$\psi(p) = \int \mathrm{d}x \, \langle p|x\rangle\langle x| \, \rangle = \int \mathrm{d}x \, \frac{\mathrm{e}^{-\mathrm{i}px/\hbar}}{\sqrt{2\pi\hbar}}\psi(x)\,. \tag{4.6.12}$$

Mathematically speaking, these are Fourier* transformations and the pair of relations states Fourier's theorem about his transformation and its inverse.

We can do what Fourier could not, namely express his theorem compactly, with the aid of Dirac's delta function. For, if we consider the iteration of (4.6.12),

$$\begin{aligned}\psi(x) &= \int \mathrm{d}p \, \frac{\mathrm{e}^{\mathrm{i}xp/\hbar}}{\sqrt{2\pi\hbar}} \int \mathrm{d}x' \, \frac{\mathrm{e}^{-\mathrm{i}px'/\hbar}}{\sqrt{2\pi\hbar}}\psi(x') \\ &= \int \mathrm{d}x' \left(\int \mathrm{d}p \, \frac{\mathrm{e}^{\mathrm{i}p(x\,-\,x')/\hbar}}{2\pi\hbar} \right) \psi(x')\,,\end{aligned} \tag{4.6.13}$$

we read off that

$$\int \mathrm{d}p \, \frac{\mathrm{e}^{\mathrm{i}p(x\,-\,x')/\hbar}}{2\pi\hbar} = \delta(x - x')\,, \tag{4.6.14}$$

the Fourier representation of the delta function — that we have implicitly used already in (4.5.11) with (4.5.14) — which is in fact the basis of the construction (4.1.18).

Given that

$$\langle x|P| \, \rangle = \frac{\hbar}{\mathrm{i}} \frac{\partial}{\partial x} \langle x| \, \rangle = \frac{\hbar}{\mathrm{i}} \frac{\partial}{\partial x} \psi(x)\,, \tag{4.6.15}$$

we can evaluate the expectation value of P in terms of the position wave function, either as

$$\langle P \rangle = \int \mathrm{d}x \, \langle \, |x\rangle\langle x|P| \, \rangle = \int \mathrm{d}x \, \psi(x)^* \frac{\hbar}{\mathrm{i}} \frac{\partial}{\partial x} \psi(x) \tag{4.6.16}$$

or as

$$\langle P \rangle = \int \mathrm{d}x \, \langle \, |P|x\rangle\langle x| \, \rangle = \int \mathrm{d}x \left(\mathrm{i}\hbar \frac{\partial}{\partial x} \psi(x)^* \right) \psi(x)\,. \tag{4.6.17}$$

*Jean Baptiste Joseph Fourier (1768–1830)

These must, of course, give the same result, and indeed they do because their difference

$$\begin{aligned}
&\int \mathrm{d}x \left(\psi(x)^* \frac{\hbar}{\mathrm{i}} \frac{\partial}{\partial x} \psi(x) - \left(\mathrm{i}\hbar \frac{\partial}{\partial x} \psi(x)^* \right) \psi(x) \right) \\
&= \frac{\hbar}{\mathrm{i}} \int \mathrm{d}x \, \frac{\partial}{\partial x} \big(\psi(x)^* \psi(x) \big) \\
&= \left(\frac{\hbar}{\mathrm{i}} \big| \psi(x) \big|^2 \right) \Bigg|_{x=-\infty}^{\infty} = 0
\end{aligned} \tag{4.6.18}$$

vanishes since $\big|\psi(x)\big|^2 \to 0$ for $x \to \pm\infty$, as is ensured by

$$\int \mathrm{d}x \, \big|\psi(x)\big|^2 = 1 \,, \tag{4.6.19}$$

but pay attention to Exercise 79.

4.7 Uncertainty relation

For any hermitian A, be it a function of X, or of P, or of both (or, possibly, involving further degrees of freedom), we have its expectation value $\langle A \rangle$ so that the difference $A - \langle A \rangle$ is an operator with vanishing expectation value. A measurement of $A - \langle A \rangle$ gives results that are positive or negative, such that their weighted sum, weighted by the relative frequency of occurrence, is zero. As a mathematical statement,

$$\Big\langle \big(A - \langle A \rangle \big) \Big\rangle = 0 \,, \tag{4.7.1}$$

it borders on a banality. Yet, how this zero average comes about can be very different. There can be lots of measurement results of A close to $\langle A \rangle$ and scattered about this value, or lots of results quite different from $\langle A \rangle$, some much larger, some much smaller. To get a quantitative handle on such differences in the data scatter, we evaluate the expectation value of $\big(A - \langle A \rangle \big)^2$. It is the mean value of positive numbers (or, at least, not negative ones) and is therefore itself positive. This permits to identify its square root, δA, with the *spread* of the measurement results of $A - \langle A \rangle$ or in fact the spread of A itself. Thus, we write

$$\delta A = \sqrt{\Big\langle \big(A - \langle A \rangle \big)^2 \Big\rangle} \tag{4.7.2}$$

as the technical definition of δA. In view of the identity

$$\begin{aligned}\Big\langle \big(A-\langle A\rangle\big)^2\Big\rangle &= \Big\langle \big(A^2-2\,\langle A\rangle\, A+\langle A\rangle^{\,2}\big)\Big\rangle \\ &= \big\langle A^2\big\rangle - 2\big\langle A\big\rangle\big\langle A\big\rangle + \big\langle A\big\rangle^2 \\ &= \big\langle A^2\big\rangle - \big\langle A\big\rangle^2 , \end{aligned} \tag{4.7.3}$$

it is clear that

$$\delta A = \sqrt{\big\langle A^2\big\rangle - \big\langle A\big\rangle^2} \tag{4.7.4}$$

is an equivalent expression that could also be used as the definition of δA. Incidentally, we note that $(\delta A)^2$, equal to both the initial and the final expression in (4.7.3), is called the *variance* of A, that is, $(\text{spread})^2 = \text{variance}$.

Now, consider two observables specified by operators A and B, along with their expectation values, $\langle A\rangle, \langle B\rangle$ and their spreads $\delta A, \delta B$. Both A and B are, of course, meant to be hermitian in the present context — otherwise $\big(A-\langle A\rangle\big)^2$ would not be assuredly positive — and so the auxiliary operator

$$C = \delta B\,(A-\langle A\rangle) + \mathrm{i}\,\delta A\,(B-\langle B\rangle) \tag{4.7.5}$$

is not hermitian, it differs from its adjoint

$$C^\dagger = \delta B\,(A-\langle A\rangle) - \mathrm{i}\,\delta A\,(B-\langle B\rangle) \tag{4.7.6}$$

by a change of sign for the second summand. Yet, the products

$$C^\dagger C \quad \text{and} \quad CC^\dagger \tag{4.7.7}$$

are both positive operators in the sense that their expectation values,

$$\begin{aligned}\Big\langle C^\dagger C\Big\rangle &= \mathrm{tr}\big(\rho\, C^\dagger C\big) \geq 0\,, \\ \Big\langle CC^\dagger\Big\rangle &= \mathrm{tr}\big(\rho\, CC^\dagger\big) \geq 0\,, \end{aligned} \tag{4.7.8}$$

are positive for *all* statistical operators ρ. This is easy to see as soon as we represent ρ as one of its many blends,

$$\rho = \sum_k \big|k\big\rangle w_k \big\langle k\big|\,, \qquad w_k > 0\,, \qquad \sum_k w_k = 1\,, \tag{4.7.9}$$

where the kets $|k\rangle$ are normalized, $\langle k|k\rangle = 1$, but otherwise arbitrary, and the weights w_k are only restricted by the stated positivity and normalization

to unit sum. Then, for example,

$$\Big\langle C^\dagger C\Big\rangle = \sum_k \underbrace{w_k}_{>0} \underbrace{\langle k|C^\dagger C|k\rangle}_{=(\langle k|C^\dagger)(C|k\rangle)=(C|k\rangle)^\dagger(C|k\rangle)\geq 0} \geq 0 \tag{4.7.10}$$

indeed.

Writing it out,

$$\begin{aligned} C^\dagger C &= (\delta B)^2\big(A-\langle A\rangle\big)^2 + (\delta A)^2\big(B-\langle B\rangle\big)^2 \\ &\quad + \mathrm{i}\,\delta A\,\delta B\Big(\big(A-\langle A\rangle\big)\big(B-\langle B\rangle\big) - \big(B-\langle B\rangle\big)\big(A-\langle A\rangle\big)\Big) \\ &= (\delta B)^2\big(A-\langle A\rangle\big)^2 + (\delta A)^2\big(B-\langle B\rangle\big)^2 \\ &\quad + \mathrm{i}\,\delta A\,\delta B\,[A,B]\,, \end{aligned} \tag{4.7.11}$$

we have

$$0 \leq \Big\langle C^\dagger C\Big\rangle = 2(\delta A\,\delta B)^2 + \delta A\,\delta B\,\big\langle \mathrm{i}[A,B]\big\rangle \tag{4.7.12}$$

and likewise

$$0 \leq \Big\langle CC^\dagger\Big\rangle = 2(\delta A\,\delta B)^2 - \delta A\,\delta B\,\big\langle \mathrm{i}[A,B]\big\rangle \tag{4.7.13}$$

or

$$\delta A\,\delta B \geq -\frac{1}{2}\big\langle \mathrm{i}[A,B]\big\rangle \quad \text{and} \quad \delta A\,\delta B \geq \frac{1}{2}\big\langle \mathrm{i}[A,B]\big\rangle \tag{4.7.14}$$

after dividing by $2\,\delta A\,\delta B$; see Exercise 83 for the case of $\delta A\,\delta B = 0$. Accordingly,

$$\delta A\,\delta B \geq \frac{1}{2}\Big|\big\langle \mathrm{i}[A,B]\big\rangle\Big|\,. \tag{4.7.15}$$

This is Robertson's* general version of Heisenberg's uncertainty relation (but see Exercises 90–92). The original version is obtained by specializing to $A = X$ and $B = P$, for which

$$\mathrm{i}[A,B] = \mathrm{i}\underbrace{[X,P]}_{=\mathrm{i}\hbar} = -\hbar \tag{4.7.16}$$

so that

$$\delta X\,\delta P \geq \frac{1}{2}\hbar\,. \tag{4.7.17}$$

*Howard Percy ROBERTSON (1903–1961)

This is Heisenberg's celebrated *uncertainty relation*, first stated in this form by Kennard.* It is arguably the most famous of all the famous easily stated results of quantum mechanics. It is sometimes, and a bit misleadingly, referred to as the "uncertainty principle" although it is not a physical principle of any kind; it is a relation, an inequality.

What is the physical content of Heisenberg's uncertainty relation for position and momentum? Let us first consider a situation in which the momentum spread δP is very small. Then the momentum is well defined, we can predict with virtual certainty the outcome of a momentum measurement. But this also means that the de Broglie wavelength $\lambda = 2\pi\hbar/p$ of (4.5.16) is a very well-defined quantity. Now, to determine the wavelength of any wave with high precision, you need a very long wave train, one that displays many regular oscillations. Such a wave train is spread out over a large area and thus there is no precision at all when you assign a position to this wave train. Accordingly, the outcome of a position measurement will be utterly unpredictable so that δX has to be large if δP is small.

The other extreme of a small δX, that is, of a well-defined position, in the sense that we can predict the outcome of a position measurement with great confidence, would in turn require a superposition of de Broglie waves with a very wide spread in wavelength and thus momentum, in order to ensure destructive interference almost everywhere, except for the small x region where the atom is located. Thus, a small δX implies a large δP, indeed.

4.8 State of minimum uncertainty

The derivation of the general uncertainty relation (4.7.15) shows that the equal sign can only hold for states described by kets $|\ \rangle$ that are eigenkets of either C or $C^\dagger$, with eigenvalue zero,

$$\text{either} \quad C\big|\ \big\rangle = 0 \quad \text{or} \quad C^\dagger\big|\ \big\rangle = 0\,. \tag{4.8.1}$$

Let us see if we can find such a minimum-uncertainty state for $A = X$ and $B = P$. It is clearly enough to consider the case $\langle X\rangle = 0$, $\langle P\rangle = 0$, and so we are looking for a ket that obeys

$$\big(\delta P\, X + \mathrm{i}\,\delta X\, P\big)\big|\ \big\rangle = 0 \tag{4.8.2}$$

with $\delta X\,\delta P = \dfrac{1}{2}\hbar$; for "$-$i" rather than "+i" see after (4.8.19).

*Earle Hesse Kennard (1885–1968)

Put bra $\langle x|$ next to this equation to turn it into a differential equation for the wave function $\psi(x) = \langle x|\,\rangle$,

$$0 = \big\langle x\big|(\delta P\, X + \mathrm{i}\,\delta X\, P)\big|\,\big\rangle = \Big(\underbrace{\delta P}_{=\hbar/(2\delta X)} x + \mathrm{i}\,\delta X \frac{\hbar}{\mathrm{i}}\frac{\partial}{\partial x}\Big)\underbrace{\langle x|\,\rangle}_{=\psi(x)} \tag{4.8.3}$$

or

$$\frac{\partial}{\partial x}\psi(x) = -\frac{x}{2(\delta X)^2}\psi(x)\,, \tag{4.8.4}$$

which is solved by

$$\psi(x) = \psi_0\, \mathrm{e}^{-\left(\frac{1}{2}x/\delta X\right)^2}\,. \tag{4.8.5}$$

The multiplicative constant ψ_0 is fixed partly by the normalization to unit integral,

$$\int \mathrm{d}x\, \big|\psi(x)\big|^2 = 1\,, \tag{4.8.6}$$

and partly by convention, that is, the freedom to choose an arbitrary overall phase factor, which we exploit by insisting on $\psi_0 > 0$. Then, the normalization says

$$\big|\psi_0\big|^2 \int \mathrm{d}x\; \mathrm{e}^{-\frac{1}{2}(x/\delta X)^2} = 1\,, \tag{4.8.7}$$

where we meet a gaussian* integral, a special case of the integral in Exercise 74, so that

$$\big|\psi_0\big|^2 \sqrt{\frac{\pi}{\frac{1}{2}/(\delta X)^2}} = \big|\psi_0\big|^2 \sqrt{2\pi}\delta X = 1 \tag{4.8.8}$$

or

$$\psi_0 = \frac{(2\pi)^{-\frac{1}{4}}}{\sqrt{\delta X}} \quad \text{for} \quad \psi_0 > 0\,. \tag{4.8.9}$$

In summary, the quest for a minimum-uncertainty state is answered affirmatively by

$$\psi(x) = \big\langle x|\,\big\rangle = \frac{(2\pi)^{-\frac{1}{4}}}{\sqrt{\delta X}}\, \mathrm{e}^{-\left(\frac{1}{2}x/\delta X\right)^2}\,. \tag{4.8.10}$$

*Karl Friedrich GAUSS (1777–1855)

The corresponding momentum wave function is also available,

$$\psi(p) = \langle p| \,\rangle = \int \mathrm{d}x \, \langle p|x\rangle\langle x| \,\rangle \tag{4.8.11}$$

or

$$\psi(p) = \frac{1}{\sqrt{2\pi\hbar}} \frac{(2\pi)^{-\frac{1}{4}}}{\sqrt{\delta X}} \int \mathrm{d}x \; \mathrm{e}^{-\left(\frac{1}{2}x/\delta X\right)^2 \,-\, \mathrm{i}px/\hbar} \,. \tag{4.8.12}$$

This is another gaussian integral, now with the parameters in Exercise 74 identified as

$$a = \frac{1}{(2\,\delta X)^2} \,, \qquad b = -\frac{\mathrm{i}}{2}\frac{p}{\hbar} \tag{4.8.13}$$

so that

$$\begin{aligned} \psi(p) &= \frac{1}{\sqrt{2\pi\hbar}} \frac{(2\pi)^{-\frac{1}{4}}}{\sqrt{\delta X}} \sqrt{\pi}\, 2\,\delta X \, \mathrm{e}^{-(p\,\delta X/\hbar)^2} \\ &= \frac{(2\pi)^{-\frac{1}{4}}}{\sqrt{\frac{\hbar}{2\delta X}}} \, \mathrm{e}^{-(p\,\delta X/\hbar)^2} \end{aligned} \tag{4.8.14}$$

or, with $\delta P = \dfrac{\hbar}{2}/\delta X$,

$$\psi(p) = \frac{(2\pi)^{-\frac{1}{4}}}{\sqrt{\delta P}} \, \mathrm{e}^{-\left(\frac{1}{2}p/\delta P\right)^2} \,. \tag{4.8.15}$$

Note how the replacements $x \leftrightarrow p$, $\delta X \leftrightarrow \delta P$ turn $\psi(x)$ and $\psi(p)$ into each other.

We should verify, as a check and a protection against silly errors, that the position and momentum uncertainties for these wave functions are indeed equal to δX and δP. Since $\langle X\rangle = 0$ and $\langle P\rangle = 0$ for obvious reasons of symmetry, we are called to verify that

$$\left\langle X^2 \right\rangle = (\delta X)^2 \,, \qquad \left\langle P^2 \right\rangle = (\delta P)^2 \,, \tag{4.8.16}$$

where it suffices to check one of them. We take $\left\langle X^2 \right\rangle$,

$$\left\langle X^2 \right\rangle = \int \mathrm{d}x \, x^2 \left|\psi(x)\right|^2 = \frac{1}{\sqrt{2\pi}} \frac{1}{\delta X} \int \mathrm{d}x \, x^2 \, \mathrm{e}^{-\frac{1}{2}(x/\delta X)^2} \,. \tag{4.8.17}$$

This integral is also a gaussian integral of some sort,

$$\int \mathrm{d}x \, x^2 \, \mathrm{e}^{-ax^2} = -\frac{\partial}{\partial a} \int \mathrm{d}x \; \mathrm{e}^{-ax^2} = -\frac{\partial}{\partial a} \sqrt{\frac{\pi}{a}} = \frac{1}{2}\sqrt{\frac{\pi}{a^3}} \tag{4.8.18}$$

so that

$$\left\langle X^2 \right\rangle = \frac{1}{\sqrt{2\pi}} \frac{1}{\delta X} \frac{1}{2} \sqrt{\frac{\pi}{(\frac{1}{2}/(\delta X)^2)^3}} = (\delta X)^2 \,, \tag{4.8.19}$$

indeed.

Finally, we remark on the apparent option to choose the other sign in (4.8.3). It would lead us first to

$$\frac{\partial}{\partial x}\psi(x) = \frac{x}{2(\delta X)^2}\psi(x) \tag{4.8.20}$$

and then to

$$\psi(x) = \psi_0 \, \mathrm{e}^{\left(\frac{1}{2}x/\delta X\right)^2} \,, \tag{4.8.21}$$

for which there is no way of satisfying the normalization condition in (4.8.6). As a consequence, this other sign choice is not an actual option.

4.9 Time dependence

So far, all remarks about position X and momentum P referred to one single instant in time. But, of course, atoms move, and so X is naturally dependent on time or, put differently, we can measure the atom's position at various instants and need position operators for all times, $X(t)$. Likewise, the momentum operator depends on time, $P(t)$.

Then, also their eigenkets and eigenbras acquire a time dependence, as illustrated by the eigenvalue equation

$$\big\langle x,t\big| \, X(t) = x \, \big\langle x,t\big| \,, \tag{4.9.1}$$

which is an equation of the same kind as (3.1.9). Small time increments δt are generated by the Hamilton operator H and small position increments δx by the momentum operator P so that a change of both gives

$$\delta\big\langle x,t\big| = \frac{\mathrm{i}}{\hbar}\big\langle x,t\big|\big(P(t)\,\delta x - H(t)\,\delta t\big) \,, \tag{4.9.2}$$

which encompasses the Schrödinger equation

$$\mathrm{i}\hbar\frac{\partial}{\partial t}\big\langle \ldots,t\big| = \big\langle \ldots,t\big|H(t) \tag{4.9.3}$$

as well as

$$\frac{\hbar}{\mathrm{i}}\frac{\partial}{\partial x}\big\langle x,\ldots,t\big| = \big\langle x,\ldots,t\big|P(t) \,, \tag{4.9.4}$$

where the ellipses stand for other, currently irrelevant labels.

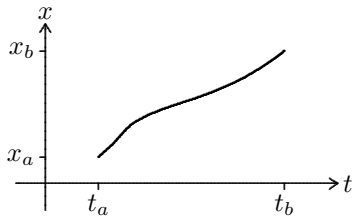
The relative sign in $P\,\delta x - H\,\delta t$ originates in the sign convention in (4.3.10), where we wrote $e^{\mathrm{i}x'P/\hbar}$ rather than $\mathrm{e}^{-\mathrm{i}x'P/\hbar}$, a choice that was not given any justification then. We shall close this gap now, with a little excursion into classical mechanics.

4.10 Excursion into classical mechanics

The question is, of course, not about the chosen sign of i in all our equations; we could just as well replace i by $-\mathrm{i}$ everywhere without the slightest consequence (in the early days of quantum mechanics, there was a temporary coexistence of both sign conventions, but nowadays, everybody uses the same convention, the one that has $[X, P] = \mathrm{i}\hbar$, rather than $-\mathrm{i}\hbar$). The question is about the relative minus sign between $P\,\delta x$ and $H\,\delta t$. And that comes straight from classical mechanics and stays with us in quantum mechanics if we wish, as we do, to maintain as good a resemblance as we can. After all, we borrow classical terminology, such as "Hamilton function" and "momentum," from classical mechanics and should, therefore, do this borrowing consistently.

We recall Lagrange's* variational principle. It says that the actual trajectory is the one for which the action

$$W_{ab} = \int_a^b \mathrm{d}t\, L \tag{4.10.1}$$



is stationary, provided the endpoints are not varied:

$$\delta W_{ab} = \delta \int_a^b \mathrm{d}t\, L = 0 \quad \text{if} \quad \delta t_a = \delta t_b = 0 \quad \text{and} \quad \delta x_a = \delta x_b = 0\,. \tag{4.10.2}$$

With $L = L(x, \dot{x}, t)$, where $\dot{x} = \dfrac{\mathrm{d}x}{\mathrm{d}t}$ is the velocity, this implies the familiar Euler–Lagrange equation

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{x}} - \frac{\partial L}{\partial x} = 0\,. \tag{4.10.3}$$

*Joseph Louis de Lagrange (1736–1813)

For the standard form of the Lagrange function L,

$$L = \frac{1}{2}M\dot{x}^2 - V(x)\,, \tag{4.10.4}$$

where M is the mass of the particle, (4.10.4) reads more explicitly

$$\frac{\mathrm{d}}{\mathrm{d}t}(M\dot{x}) + \frac{\partial V}{\partial x} = 0 \quad \text{or} \quad M\frac{\mathrm{d}^2 x}{\mathrm{d}t^2} = F\,, \tag{4.10.5}$$

with the force $F = -\dfrac{\partial V}{\partial x}$. This is the Newton's familiar equation of motion.

Now, if we also allow endpoint variations, the variational principle says that all first-order changes come from the endpoints,

$$\delta W_{ab} = \delta \int_a^b \mathrm{d}t\, L = G_b - G_a\,, \tag{4.10.6}$$

where G_a involves the infinitesimal endpoint variations δx_a and δt_a of the initial position and initial time and G_b involves the variations δx_b and δt_b of the final position and the final time. To identify these *generators* G_a, G_b, we introduce a path parameter τ, $0 \le \tau \le 1$, and regard the trajectory in x and t as a function of τ,

$$t = t(\tau)\,, \quad x = x(\tau) \quad \text{with} \quad \begin{cases} t = t_a\,,\ x = x_a \text{ for } \tau = 0\,, \\ t = t_b\,,\ x = x_b \text{ for } \tau = 1\,. \end{cases} \tag{4.10.7}$$

Then,

$$\dot{x} = \frac{\mathrm{d}x/\mathrm{d}\tau}{\mathrm{d}t/\mathrm{d}\tau} \equiv \frac{x'}{t'} \quad \text{and} \quad \mathrm{d}t = \mathrm{d}\tau\,\frac{\mathrm{d}t}{\mathrm{d}\tau} = \mathrm{d}\tau\, t'\,, \tag{4.10.8}$$

and we write

$$W_{ab} = \int_0^1 \mathrm{d}\tau\ t' L(x, \dot{x}, t). \tag{4.10.9}$$

Now, vary $t(\tau)$ and $x(\tau)$:

$$\begin{aligned} &t(\tau) \to t(\tau) + \delta t(\tau)\,, \\ &x(\tau) \to x(\tau) + \delta x(\tau)\,, \\ &t'(\tau) \to t'(\tau) + \delta t'(\tau) = t'(\tau) + \frac{\mathrm{d}}{\mathrm{d}\tau}\delta t(\tau) \end{aligned} \tag{4.10.10}$$

to get

$$\delta W_{ab} = \int_0^1 \mathrm{d}\tau \left(\delta t' L + t'\delta x \frac{\partial L}{\partial x} + t'\delta\dot{x}\frac{\partial L}{\partial \dot{x}} + t'\delta t\,, \frac{\partial L}{\partial t} \right), \tag{4.10.11}$$

where

$$\delta\dot{x} = \delta\frac{x'}{t'} = \frac{\delta x'}{t'} - \frac{x'\delta t'}{t'^2} = \frac{1}{t'}\big(\delta x' - \delta t'\,\dot{x}\big)\,. \tag{4.10.12}$$

With

$$\delta t' L = \frac{\mathrm{d}}{\mathrm{d}\tau}(\delta t\, L) - \delta t \frac{\mathrm{d}}{\mathrm{d}\tau} L \tag{4.10.13}$$

and

$$\begin{aligned} t'\delta\dot{x}\frac{\partial L}{\partial \dot{x}} &= \delta x' \frac{\partial L}{\partial \dot{x}} - \delta t'\,\dot{x}\frac{\partial L}{\partial \dot{x}} \\ &= \frac{\mathrm{d}}{\mathrm{d}\tau}\left(\delta x \frac{\partial L}{\partial \dot{x}} - \delta t\,\dot{x}\frac{\partial L}{\partial \dot{x}}\right) \\ &\quad - \delta x \frac{\mathrm{d}}{\mathrm{d}\tau}\frac{\partial L}{\partial \dot{x}} + \delta t \frac{\mathrm{d}}{\mathrm{d}\tau}\left(\dot{x}\frac{\partial L}{\partial \dot{x}}\right), \end{aligned} \tag{4.10.14}$$

we arrive at

$$\begin{aligned} \delta W_{ab} = \int_0^1 \mathrm{d}\tau \Bigg(& \frac{\mathrm{d}}{\mathrm{d}\tau}\left(\delta x \frac{\partial L}{\partial \dot{x}} - \delta t\left(\dot{x}\frac{\partial L}{\partial \dot{x}} - L\right)\right) \\ & + \delta t\left(-\frac{\mathrm{d}}{\mathrm{d}\tau} L + t'\frac{\partial L}{\partial t} + \frac{\mathrm{d}}{\mathrm{d}\tau}\left(\dot{x}\frac{\partial L}{\partial \dot{x}}\right)\right) \\ & + \delta x\left(t'\frac{\partial L}{\partial x} - \frac{\mathrm{d}}{\mathrm{d}\tau}\frac{\partial L}{\partial \dot{x}}\right)\Bigg)\,. \end{aligned} \tag{4.10.15}$$

The total τ derivative gives the endpoint contributions

$$\delta W_{ab} = \left(\delta x \frac{\partial L}{\partial \dot{x}} - \delta t\left(\dot{x}\frac{\partial L}{\partial \dot{x}} - L\right)\right)\Bigg|_a^b = G_b - G_a \tag{4.10.16}$$

with

$$G = \delta x\,\frac{\partial L}{\partial \dot{x}} - \delta t\left(\dot{x}\frac{\partial L}{\partial \dot{x}} - L\right) = \delta x\, p - \delta t\, H\,, \tag{4.10.17}$$

which identifies the *momentum* $p = \dfrac{\partial L}{\partial \dot{x}}$ and the energy, or the *Hamilton function*,

$$H = \dot{x}\frac{\partial L}{\partial \dot{x}} - L\,. \tag{4.10.18}$$

The internal contributions that are proportional to the independent infinitesimal path variations δx and δt must vanish so that the δx term implies

$$t' \frac{\partial L}{\partial x} - \frac{\mathrm{d}}{\mathrm{d}\tau} \frac{\partial L}{\partial \dot{x}} = 0 \,. \tag{4.10.19}$$

With $\dfrac{\mathrm{d}}{\mathrm{d}\tau} = t' \dfrac{\mathrm{d}}{\mathrm{d}t}$, this is the Euler–Lagrange equation of motion (4.10.3), as it should be.

The generator of (4.10.17),

$$G = \delta x \, p - \delta t \, H \,, \tag{4.10.20}$$

provides $p = \dfrac{\partial L}{\partial \dot{x}}$ and $H = \dot{x} \dfrac{\partial L}{\partial \dot{x}} - L$, which are just the usual kinetic momentum

$$p = M \dot{x} \tag{4.10.21}$$

and the usual energy

$$H = \frac{1}{2} M \dot{x}^2 + V(x) \tag{4.10.22}$$

for the Lagrange function (4.10.4).

We recognize in (4.10.20) the same formal structure as the one in (4.9.2), only that we are talking about the *momentum operator* and the *Hamilton operator* there, with the same fundamental minus sign between them. Note further that the response of the Hamilton function to arbitrary variations,

$$\begin{aligned} \delta H &= \delta \dot{x} \, p + \dot{x} \delta p - \delta x \frac{\partial L}{\partial x} - \delta \dot{x} \underbrace{\frac{\partial L}{\partial \dot{x}}}_{= p} - \delta t \frac{\partial L}{\partial t} \\ &= -\delta x \frac{\partial L}{\partial x} + \dot{x} \, \delta p - \delta t \frac{\partial L}{\partial t} \,, \end{aligned} \tag{4.10.23}$$

shows that the natural variables of H are x, p, and t,

$$H = H(x, p, t), \tag{4.10.24}$$

whereas the natural variables of L are x, $\dot{x}$, and t. We must therefore express H as a function of x, p, and t, which results in

$$H = \frac{1}{2M} p^2 + V(x) \tag{4.10.25}$$

for the L in (4.10.4).

4.11 Hamilton operator, Schrödinger equation

We borrow this structure from classical mechanics and conjecture that the Hamilton operator of an atom of mass M, in motion under the influence of the force that is associated with the potential energy V, is

$$H = \frac{1}{2M}P^2 + V(X)\,. \tag{4.11.1}$$

In particular, for force-free motion ("free particle"), we only have the kinetic energy term,

$$H = \frac{1}{2M}P^2\,. \tag{4.11.2}$$

The Schrödinger equation then reads

$$\mathrm{i}\hbar\frac{\partial}{\partial t}\big\langle \ldots, t\big| = \big\langle \ldots, t\big|\frac{1}{2M}P^2 \tag{4.11.3}$$

for a general bra. Depending on the quantum numbers the ellipsis refers to, it becomes a numerical statement about time-dependent wave functions, with a large variety of appearances. For position wave functions, for example, we get

$$\mathrm{i}\hbar\frac{\partial}{\partial t}\psi(x,t) = -\frac{\hbar^2}{2M}\frac{\partial^2}{\partial x^2}\psi(x,t) \quad \text{with} \quad \psi(x,t) = \big\langle x,t\big|\,\big\rangle\,, \tag{4.11.4}$$

whereas

$$\mathrm{i}\hbar\frac{\partial}{\partial t}\psi(p,t) = \frac{p^2}{2M}\psi(p,t) \quad \text{with} \quad \psi(p,t) = \big\langle p,t\big|\,\big\rangle \tag{4.11.5}$$

applies to momentum wave functions. They need to be solved subject to initial conditions that specify the wave functions at $t = 0$, say,

$$\big\langle x,t=0\big|\,\big\rangle = \psi_0(x) \quad \text{or} \quad \big\langle p,t=0\big|\,\big\rangle = \psi_0(p)\,. \tag{4.11.6}$$

Since we can always switch from position wave functions to momentum wave functions and *vice versa*, solving one of the equations is tantamount to solving the other.

Before proceeding, let us recognize that these differential equations for the evolution of the wave functions for force-free motion are just particular cases of a much more general set of equations, namely the differential equations for arbitrary Hamilton operators

$$H(t) = H\big(X(t), P(t), t\big) \tag{4.11.7}$$

that have a dynamical time dependence because they are functions of the dynamical variables $P(t)$ and $X(t)$, and may have an additional parametric time dependence, if the functional dependence on P and X changes in time. Most typical is the form

$$H = \frac{1}{2M}P^2 + V(X,t)\,, \tag{4.11.8}$$

the sum of the standard kinetic energy $P^2/(2M)$ and the situation-dependent potential energy $V(X,t)$ that might possess a parametric t dependence.

Whatever the actual form of H as a function of P, X, and t, we get

$$\begin{aligned} \mathrm{i}\hbar\frac{\partial}{\partial t}\psi(x,t) = \mathrm{i}\hbar\frac{\partial}{\partial t}\big\langle x,t\big|\,\big\rangle &= \big\langle x,t\big|H\big(\underset{\downarrow}{X}(t)\,,\underset{\downarrow}{P}(t)\,,t\big)\big|\,\big\rangle \\ &\qquad\qquad\quad x \quad \frac{\hbar}{\mathrm{i}}\frac{\partial}{\partial x} \\ &= H\bigg(x,\frac{\hbar}{\mathrm{i}}\frac{\partial}{\partial x},t\bigg)\underbrace{\big\langle x,t\big|\,\big\rangle}_{=\psi(x,t)} \end{aligned} \tag{4.11.9}$$

or

$$\mathrm{i}\hbar\frac{\partial}{\partial t}\psi(x,t) = H\bigg(x,\frac{\hbar}{\mathrm{i}}\frac{\partial}{\partial x},t\bigg)\psi(x,t) \tag{4.11.10}$$

for the position wave function $\psi(x,t) = \langle x,t|\,\rangle$. This differential equation, formally obtained by replacing $X \to x$ and $P \to \dfrac{\hbar}{\mathrm{i}}\dfrac{\partial}{\partial x}$ in $H(X,P,t)$ to turn the Hamilton operator into a differential operator, is frequently referred to as *the* Schrödinger equation. It is, in fact, the numerical version of the Schrödinger equation that is most frequently used in practice, but nonetheless it is only one of many Schrödinger equations. In particular, we have

$$\begin{aligned} \mathrm{i}\hbar\frac{\partial}{\partial t}\psi(p,t) = \mathrm{i}\hbar\frac{\partial}{\partial t}\big\langle p,t\big|\,\big\rangle &= \big\langle p,t\big|H\big(\underset{\downarrow}{X}(t)\,,\underset{\downarrow}{P}(t)\,,t\big)\big|\,\big\rangle \\ &\qquad\qquad\quad \mathrm{i}\hbar\frac{\partial}{\partial p} \quad p \\ &= H\bigg(\mathrm{i}\hbar\frac{\partial}{\partial p},p,t\bigg)\underbrace{\big\langle p,t\big|\,\big\rangle}_{=\psi(p,t)} \end{aligned} \tag{4.11.11}$$

or

$$\mathrm{i}\hbar\frac{\partial}{\partial t}\psi(p,t) = H\left(\mathrm{i}\hbar\frac{\partial}{\partial p}, p, t\right)\psi(p,t) \tag{4.11.12}$$

for the momentum wave function $\psi(p,t) = \langle p,t|\,\rangle$. And there are many more Schrödinger equations for other choices of wave functions.

The popularity of the Schrödinger equation for $\psi(x,t)$ originates in the circumstance that, usually, the momentum dependence of the Hamilton operator is quite simple, such as the one in (4.11.8), whereas the position dependence can be rather complicated. For Hamilton operators of this standard structure, we get

$$\mathrm{i}\hbar\frac{\partial}{\partial t}\psi(x,t) = \left(-\frac{\hbar^2}{2M}\frac{\partial^2}{\partial x^2} + V(x,t)\right)\psi(x,t)\,. \tag{4.11.13}$$

By all counts, this must be the numerical version of the Schrödinger equation that is printed most often by far.

4.12 Time transformation function

Generally speaking, to solve such a Schrödinger equation means to express the wave function at the later time t in terms of the wave function at the initial time $t = 0$. That is, for example,

$$\begin{aligned}\psi(x,t) &= \langle x,t|\,\rangle\langle x,t|\int \mathrm{d}x'\,|x',0\rangle\underbrace{\langle x',0|\,\rangle}_{=\psi_0(x')}\\ &= \int \mathrm{d}x'\,\langle x,t|x',0\rangle\,\psi_0(x')\end{aligned} \tag{4.12.1}$$

or

$$\begin{aligned}\psi(x,t) = \langle x,t|\,\rangle &= \langle x,t|\int \mathrm{d}p\,|p,0\rangle\underbrace{\langle p,0|\,\rangle}_{=\psi_0(p)}\\ &= \int \mathrm{d}p\,\langle x,t|p,0\rangle\,\psi_0(p)\end{aligned} \tag{4.12.2}$$

if we express the position wave function at t in terms of either the position wave function at $t = 0$ or the momentum wave function at $t = 0$. In view of

$$\psi_0(x) = \int \mathrm{d}p\,\langle x,0|p,0\rangle\underbrace{\langle p,0|\,\rangle}_{=\psi_0(p)} = \int \mathrm{d}p\,\frac{\mathrm{e}^{\mathrm{i}xp/\hbar}}{\sqrt{2\pi\hbar}}\,\psi_0(p) \tag{4.12.3}$$

and its inverse relation

$$\psi_0(p) = \int \mathrm{d}x \, \frac{\mathrm{e}^{-\mathrm{i}xp/\hbar}}{\sqrt{2\pi\hbar}} \, \psi_0(x) \,, \tag{4.12.4}$$

which repeat (4.6.12), we can go from one formulation to the other quite easily. Whether we calculate the *xx time-transformation function* $\langle x,t|x',0\rangle$ or the *xp time-transformation function* $\big\langle x,t\big|p,0\big\rangle$ or perhaps another one does not matter because we can turn one into the other, as illustrated by

$$\begin{aligned} \big\langle x,t\big|x',0\big\rangle &= \int \mathrm{d}p \, \big\langle x,t\big|p,0\big\rangle\big\langle p,0\big|x',0\big\rangle \\ &= \int \mathrm{d}p \, \big\langle x,t\big|p,0\big\rangle \frac{\mathrm{e}^{-\mathrm{i}px'/\hbar}}{\sqrt{2\pi\hbar}} \,, \end{aligned} \tag{4.12.5}$$

which tells us how to obtain $\langle x,t|x',0\rangle$ when we know $\langle x,t|p,0\rangle$.

Since

$$\mathrm{i}\hbar\frac{\partial}{\partial t}\big\langle x,t\big| = \big\langle x,t\big|H\big(X(t),P(t),t\big) = H\bigg(x,\frac{\hbar}{\mathrm{i}}\frac{\partial}{\partial x},t\bigg)\big\langle x,t\big| \,, \tag{4.12.6}$$

all transformation functions $\langle x,t|\ldots,0\rangle$ obey the same differential Schrödinger equation. What distinguishes them from each other are the corresponding initial conditions. For $\langle x,t|x',0\rangle$, we have a delta function,

$$\big\langle x,t\big|x',0\big\rangle \to \delta(x-x') \quad \text{as} \quad t\to 0 \,, \tag{4.12.7}$$

whereas we have a plane wave for $\langle x,t|p,0\rangle$,

$$\big\langle x,t\big|p,0\big\rangle \to \frac{\mathrm{e}^{\mathrm{i}xp/\hbar}}{\sqrt{2\pi\hbar}} \quad \text{as} \quad t\to 0 \,. \tag{4.12.8}$$

It is often a matter of sheer convenience that makes us prefer one time transformation function over the other.

Chapter 5

Elementary Examples

5.1 Force-free motion

5.1.1 *Time-transformation functions*

Now, after these remarks of a quite general nature, let us return to the problem of finding the solution to the Schrödinger equation for an atom in force-free motion,

$$\mathrm{i}\hbar\frac{\partial}{\partial t}\psi(x,t) = -\frac{\hbar^2}{2M}\frac{\partial^2}{\partial x^2}\psi(x,t)\,. \tag{5.1.1}$$

It is convenient to first consider the xp transformation function, for which

$$\mathrm{i}\hbar\frac{\partial}{\partial t}\big\langle x,t\big|p,0\big\rangle = \big\langle x,t\big|H\big|p,0\big\rangle \tag{5.1.2}$$

with

$$H = \frac{1}{2M}P(t)^2 = \frac{1}{2M}P(0)^2 \tag{5.1.3}$$

so that

$$H\big|p,0\big\rangle = \big|p,0\big\rangle\frac{p^2}{2M}\,. \tag{5.1.4}$$

Here, we recall from (3.2.9) that a Hamilton operator with no parametric time dependence does not depend on time at all, $H\big(X(t),P(t)\big) = H\big(X(0),P(0)\big)$, and appreciate the resulting simplification,

$$\mathrm{i}\hbar\frac{\partial}{\partial t}\big\langle x,t\big|p,0\big\rangle = \big\langle x,t\big|p,0\big\rangle\frac{p^2}{2M}\,. \tag{5.1.5}$$

This is an elementary differential equation, with the solution

$$\langle x,t|p,0\rangle = \frac{\mathrm{e}^{\mathrm{i}xp/\hbar}}{\sqrt{2\pi\hbar}}\,\mathrm{e}^{-\frac{\mathrm{i}}{\hbar}\frac{p^2}{2M}t}\,, \tag{5.1.6}$$

where the time-independent prefactor incorporates the initial condition (4.12.8). Accordingly, we can now get the position wave function at time t from the momentum wave function at time $t=0$,

$$\begin{aligned}\psi(x,t) &= \int \mathrm{d}p\,\langle x,t|p,0\rangle\langle p,0|\ \rangle \\ &= \int \mathrm{d}p\,\frac{\mathrm{e}^{\mathrm{i}xp/\hbar}}{\sqrt{2\pi\hbar}}\,\mathrm{e}^{-\frac{\mathrm{i}}{\hbar}\frac{p^2}{2M}t}\psi_0(p)\,.\end{aligned} \tag{5.1.7}$$

We use

$$\begin{aligned}\psi_0(p) = \langle p,0|\ \rangle &= \int \mathrm{d}x\,\langle p,0|x,0\rangle\langle x,0|\ \rangle \\ &= \int \mathrm{d}x\,\frac{\mathrm{e}^{-\mathrm{i}px/\hbar}}{\sqrt{2\pi\hbar}}\psi_0(x)\end{aligned} \tag{5.1.8}$$

to introduce the position wave function at time $t=0$,

$$\begin{aligned}\psi(x,t) &= \int \mathrm{d}p\,\frac{\mathrm{e}^{\mathrm{i}xp/\hbar}}{\sqrt{2\pi\hbar}}\,\mathrm{e}^{-\frac{\mathrm{i}}{\hbar}\frac{p^2}{2M}t}\int \mathrm{d}x'\,\frac{\mathrm{e}^{-\mathrm{i}px'/\hbar}}{\sqrt{2\pi\hbar}}\psi_0(x') \\ &= \int \mathrm{d}x'\int \mathrm{d}p\,\frac{\mathrm{e}^{\mathrm{i}(x-x')p/\hbar}}{2\pi\hbar}\,\mathrm{e}^{-\frac{\mathrm{i}}{\hbar}\frac{p^2}{2M}t}\psi_0(x')\,,\end{aligned} \tag{5.1.9}$$

which we compare with

$$\psi(x,t) = \int \mathrm{d}x'\,\langle x,t|x',0\rangle\,\psi_0(x') \tag{5.1.10}$$

to read off that

$$\langle x,t|x',0\rangle = \int \mathrm{d}p\,\frac{\mathrm{e}^{\mathrm{i}(x-x')p/\hbar}}{2\pi\hbar}\,\mathrm{e}^{-\frac{\mathrm{i}}{\hbar}\frac{p^2}{2M}t}\,. \tag{5.1.11}$$

Of course, this is nothing but the relation

$$\langle x,t|x',0\rangle = \int \mathrm{d}p\,\langle x,t|p,0\rangle\langle p,0|x',0\rangle \tag{5.1.12}$$

with the particular force-free version of $\langle x,t|p,0\rangle$ inserted along with the universal $\langle p,0|x',0\rangle = \mathrm{e}^{-\mathrm{i}px'/\hbar}/\sqrt{2\pi\hbar}$.

After completing the square in the exponent,

$$-\frac{\mathrm{i}}{\hbar}\frac{p^2}{2M}t + \frac{\mathrm{i}}{\hbar}(x-x')p = -\frac{\mathrm{i}}{\hbar}\frac{t}{2M}\left(p - \frac{M}{t}(x-x')\right)^2 + \frac{\mathrm{i}}{\hbar}\frac{M}{2t}(x-x')^2\,, \tag{5.1.13}$$

the remaining integral is of gaussian type and thus easily evaluated,

$$\begin{aligned}\big\langle x,t\big|x',0\big\rangle &= \int \mathrm{d}p\,\frac{1}{2\pi\hbar}\,\mathrm{e}^{-\frac{\mathrm{i}}{\hbar}\frac{t}{2M}\left(p-\frac{M}{t}(x-x')\right)^2}\,\mathrm{e}^{\frac{\mathrm{i}}{\hbar}\frac{M}{2t}(x-x')^2} \\ &= \frac{1}{2\pi\hbar}\sqrt{\frac{\pi}{\frac{\mathrm{i}}{\hbar}\frac{t}{2M}}}\,\mathrm{e}^{\frac{\mathrm{i}}{\hbar}\frac{M}{2t}(x-x')^2}\,,\end{aligned} \tag{5.1.14}$$

or finally

$$\big\langle x,t\big|x',0\big\rangle = \sqrt{\frac{M}{\mathrm{i}2\pi\hbar t}}\,\mathrm{e}^{\frac{\mathrm{i}}{\hbar}\frac{M}{2t}(x-x')^2}\,. \tag{5.1.15}$$

The initial condition (4.12.7) is correctly incorporated, inasmuch as we can verify the integral properly

$$\int \mathrm{d}x'\,\big\langle x,t\big|x',0\big\rangle = \sqrt{\frac{M}{\mathrm{i}2\pi\hbar t}}\sqrt{\frac{\pi}{-\frac{\mathrm{i}}{\hbar}\frac{M}{2t}}} = 1\,, \tag{5.1.16}$$

and note that $\langle x,t|x',0\rangle$ oscillates arbitrarily rapidly for $x' \neq x$ and very small t so that any x' integration with a smooth function will not have noticeable contributions beyond the immediate vicinity of $x' \cong x$, where $\langle x,t|x',0\rangle$ is very large for small t.

5.1.2 *Spreading of the wave function*

The constancy in time of the Hamilton operator $H = P^2/(2M)$ implies a constancy in time of the momentum operator P, and this must also be evident from Heisenberg's equation of motion for $P(t)$. Indeed,

$$\frac{\mathrm{d}}{\mathrm{d}t}P(t) = \frac{1}{\mathrm{i}\hbar}\big[P(t),H\big] = \frac{1}{\mathrm{i}\hbar}\left[P(t),\frac{1}{2M}P(t)^2\right] = 0\,, \tag{5.1.17}$$

and so

$$P(t) = P(0)\,. \tag{5.1.18}$$

Likewise, we have

$$\frac{\mathrm{d}}{\mathrm{d}t}X(t) = \frac{1}{\mathrm{i}\hbar}\left[X(t),\frac{1}{2M}P(t)^2\right] = \frac{1}{M}P(t)\,, \tag{5.1.19}$$

where the right-hand side — the velocity operator — is constant,

$$\frac{\mathrm{d}}{\mathrm{d}t}X(t) = \frac{1}{M}P(0)\,, \tag{5.1.20}$$

so that we can integrate immediately to arrive at

$$X(t) = X(0) + \frac{t}{M}P(0)\,. \tag{5.1.21}$$

Note that these equations look exactly the same as their classical counterparts, but now they are equations about time-dependent operators, rather than numbers.

Upon taking expectation values, we obtain numerical statements,

$$\big\langle P(t)\big\rangle = \big\langle P(0)\big\rangle\,, \qquad \big\langle X(t)\big\rangle = \big\langle X(0)\big\rangle + \frac{t}{M}\big\langle P(0)\big\rangle\,, \tag{5.1.22}$$

which are even closer counterparts to the classical analogs. Aiming at an evaluation of the position and momentum spreads, $\delta X(t)$ and $\delta P(t)$, we also find the expectation values of the squared operators,

$$\begin{aligned}
\Big\langle P(t)^2\Big\rangle &= \Big\langle P(0)^2\Big\rangle\,,\\
\Big\langle X(t)^2\Big\rangle &= \left\langle \left(X(0) + \frac{t}{M}P(0)\right)^2\right\rangle\\
&= \Big\langle X(0)^2\Big\rangle + \frac{t^2}{M^2}\Big\langle P(0)^2\Big\rangle\\
&\quad + \frac{t}{M}\Big\langle \big(X(0)P(0) + P(0)X(0)\big)\Big\rangle\,.
\end{aligned} \tag{5.1.23}$$

Accordingly, for P, we get the variance

$$\begin{aligned}
\delta P(t)^2 &= \Big\langle P(t)^2\Big\rangle - \big\langle P(t)\big\rangle^{\,2}\\
&= \Big\langle P(0)^2\Big\rangle - \big\langle P(0)\big\rangle^{\,2} = \delta P(0)^2\,,
\end{aligned} \tag{5.1.24}$$

as we anticipated — when no force is acting, there is no change in momentum whatsoever — and for the variance of X, we obtain

$$\begin{aligned}
\delta X(t)^2 &= \Big\langle X(t)^2\Big\rangle - \big\langle X(t)\big\rangle^{\,2}\\
&= \Big\langle X(0)^2\Big\rangle - \big\langle X(0)\big\rangle^{\,2} + \frac{t^2}{M^2}\Big(\Big\langle P(0)^2\Big\rangle - \big\langle P(0)\big\rangle^{\,2}\Big)\\
&\quad + \frac{t}{M}\Big(\Big\langle \big(X(0)P(0) + P(0)X(0)\big)\Big\rangle - 2\,\big\langle X(0)\big\rangle\,\big\langle P(0)\big\rangle\Big)
\end{aligned} \tag{5.1.25}$$

or

$$\delta X(t)^2 = \delta X(0)^2 + \left(\frac{t}{M}\delta P(0)\right)^2 + \frac{2t}{M}\left(\left\langle\frac{X(0)P(0)+P(0)X(0)}{2}\right\rangle - \big\langle X(0)\big\rangle\big\langle P(0)\big\rangle\right). \quad (5.1.26)$$

This tells us that the spread in position *does* depend on time, and more: it always grows linearly with t for very late times t because for large t, the dominating term is the one $\propto t^2$ so that

$$\delta X(t) \cong \frac{t}{M}\delta P(0) \qquad \text{for very late times } t. \quad (5.1.27)$$

This phenomenon is colloquially known as the "spreading of the wave function." It is, in fact, something rather natural and something rather obvious: Because we do not know the initial velocity precisely but only with an error measured by $\delta P/M$, our predictions about the object's position will become less precise in time. This is *all* there is to it. Do not imagine that the atom becomes spread out; what is spreading is the wave function by which we describe the atom, not the atom itself. When you look for it, you will always find the atom in one complete piece at a certain place, but if you wait too long before looking, you cannot guess very well where you will find the atom. The imprecision is in our capability of predicting the outcome of a position measurement; follow up with Exercises 98 and 99.

Of course, we can also see this spreading of the wave function by looking at the wave function itself rather than the expectation values that go with it. Fitting to these exercises, we take the wave function in (4.8.10) for $\psi(x, t=0) = \psi_0(x)$,

$$\psi_0(x) = \frac{(2\pi)^{-\frac{1}{4}}}{\sqrt{\delta X}}\,\mathrm{e}^{-\left(\frac{1}{2}x/\delta X\right)^2}, \quad (5.1.28)$$

and have

$$\psi(x,t) = \int \mathrm{d}x'\,\sqrt{\frac{M}{\mathrm{i}2\pi\hbar t}}\,\mathrm{e}^{\frac{\mathrm{i}}{\hbar}\frac{M}{2t}(x-x')^2}\,\frac{(2\pi)^{-\frac{1}{4}}}{\sqrt{\delta X}}\,\mathrm{e}^{-\left(\frac{1}{2}x'/\delta X\right)^2}. \quad (5.1.29)$$

The exponent is a quadratic function of x', so this is another gaussian integral, which we can evaluate explicitly without much ado. First, we complete a square, for which the identity

$$a(x'-x_1)^2 + b(x'-x_2)^2 = (a+b)\left(x' - \frac{ax_1+bx_2}{a+b}\right)^2 + \frac{ab}{a+b}(x_1-x_2)^2 \quad (5.1.30)$$

is useful. Here, we use it for

$$a = \left(\frac{1}{2\delta X}\right)^2, \quad b = -\frac{\mathrm{i}}{\hbar}\frac{M}{2t}, \quad x_1 = 0, \quad x_2 = x \tag{5.1.31}$$

and get

$$\left(\frac{x'}{2\delta X}\right)^2 - \frac{\mathrm{i}}{\hbar}\frac{M}{2t}(x-x')^2 = \left(\frac{1}{(2\delta X)^2} - \frac{\mathrm{i}}{\hbar}\frac{M}{2t}\right)\big(x' - (\text{something})\big)^2 + \frac{x^2}{(2\delta X)^2 + \mathrm{i}2\hbar t/M}\,. \tag{5.1.32}$$

We do not have to work out the "something," since it just amounts to a shift of the integration variable and has no relevance for the value of the x' integral, and we used

$$\frac{ab}{a+b} = \frac{1}{1/a + 1/b} \tag{5.1.33}$$

in the last term (this is half of the harmonic mean of a and b). After writing

$$(2\delta X)^2 + \underbrace{\mathrm{i}2\hbar}_{=4\mathrm{i}\,\delta X\,\delta P} t/M = (2\delta X)^2 \underbrace{\left(1 + \mathrm{i}\frac{t}{M}\frac{\delta P}{\delta X}\right)}_{=1/\epsilon(t)} \tag{5.1.34}$$

or

$$\frac{1}{a} + \frac{1}{b} = \frac{a+b}{ab} = \frac{1}{a\epsilon}\,, \tag{5.1.35}$$

which introduces $\epsilon(t)$, we now note that the x-independent factor that results from the gaussian integral is

$$\underbrace{\sqrt{\frac{M}{\mathrm{i}2\pi\hbar t}}}_{=\sqrt{b/\pi}} \frac{(2\pi)^{-\frac{1}{4}}}{\sqrt{\delta X}} \sqrt{\frac{\pi}{a+b}} = \frac{(2\pi)^{-\frac{1}{4}}}{\sqrt{\delta X}} \sqrt{\frac{b}{a+b}} = \frac{(2\pi)^{-\frac{1}{4}}}{\sqrt{\delta X}} \sqrt{\epsilon}\,. \tag{5.1.36}$$

Accordingly,

$$\psi(x,t) = \frac{(2\pi)^{-\frac{1}{4}}}{\sqrt{\delta X/\epsilon(t)}}\, \mathrm{e}^{-\epsilon(t)\left(\frac{1}{2}x/\delta X\right)^2}$$
$$\text{with} \quad \epsilon(t) = \left(1 + \mathrm{i}\frac{t}{M}\frac{\delta P}{\delta X}\right)^{-1}. \tag{5.1.37}$$

The wave function keeps its gaussian shape in the course of time but acquires an imaginary part. The spreading becomes particularly visible in the associated probability density $|\psi(x,t)|^2$,

$$\big|\psi(x,t)\big|^2 = \frac{1}{\sqrt{2\pi}}\frac{1}{\delta X}\big|\epsilon(t)\big|\,\mathrm{e}^{-\frac{1}{2}(x/\delta X)^2\,\mathrm{Re}\left(\epsilon(t)\right)}\,, \tag{5.1.38}$$

where we meet $\big|\epsilon(t)\big|$ and $\mathrm{Re}\big(\epsilon(t)\big)$, which are related by

$$\big|\epsilon(t)\big|^2 = \mathrm{Re}\big(\epsilon(t)\big) = \left(1+\left(\big(\frac{t\,\delta P}{M\,\delta X}\right)^2\right)^{-1}. \tag{5.1.39}$$

Therefore,

$$\big|\psi(x,t)\big|^2 = \frac{1}{\sqrt{2\pi}}\frac{1}{\delta X(t)}\,\mathrm{e}^{-\frac{1}{2}\left(x/\delta X(t)\right)^2} \tag{5.1.40}$$

with

$$\begin{aligned}\delta X(t) &= \frac{\delta X}{\big|\epsilon(t)\big|} = \frac{\delta X}{\sqrt{\mathrm{Re}\big(\epsilon(t)\big)}}\\ &= \delta X\sqrt{1+\left(\frac{t\,\delta P}{M\,\delta X}\right)^2} = \sqrt{\big(\delta X\big)^2+\left(\frac{t}{M}\delta P\right)^2}\,.\end{aligned} \tag{5.1.41}$$

Not surprisingly, this agrees with the spread $\delta X(t)$ found in Exercise 98. Upon comparing with the general expression (5.1.26), we conclude that for the chosen $\psi_0(x)$ the term $\propto t$ vanishes, that is,

$$\left\langle\frac{X(0)P(0)+P(0)X(0)}{2}\right\rangle = \big\langle X(0)\big\rangle\,\big\langle P(0)\big\rangle\,, \tag{5.1.42}$$

consistent with the observation in Exercise 99.

The spreading has yet another aspect in addition to the growing spread in position. Since $\big\langle X(t)\big\rangle = 0$ and $\big\langle P(t)\big\rangle = 0$ for all t, the atom appears to be stationary; so what is the probability that the dynamical evolution does not change anything? In other, more technical words: What is the probability of finding the atom, after the lapse of time t, in a state that has, at that later time, the wave function that the atom had initially,

$$\text{prob(no change after time } t) = \left|\int \mathrm{d}x\;\underbrace{\psi_0(x)^*}_{\text{before the spreading}\,\nearrow}\;\underbrace{\psi(x,t)}_{\nwarrow\,\text{after the spreading}}\right|^2 ? \tag{5.1.43}$$

Since we are dealing with a product of gaussian wave functions, we can evaluate the integral quite easily,

$$\int \mathrm{d}x\ \psi_0(x)^*\psi(x,t) = \int \mathrm{d}x\ \frac{(2\pi)^{-\frac{1}{4}}}{\sqrt{\delta X}}\,\mathrm{e}^{-(\frac{1}{2}x/\delta X)^2}\,\frac{(2\pi)^{-\frac{1}{4}}}{\sqrt{\delta X/\epsilon}}\,\mathrm{e}^{-\epsilon(\frac{1}{2}x/\delta X)^2}$$
$$= \frac{1}{\sqrt{2\pi/\epsilon}\,\delta X}\sqrt{\frac{\pi}{\frac{1+\epsilon}{(2\delta X)^2}}} = \sqrt{\frac{2\epsilon}{1+\epsilon}}\,, \tag{5.1.44}$$

and arrive at

$$\text{prob(no change after time } t) = \left|\frac{2\epsilon}{1+\epsilon}\right| = \sqrt{\frac{4|\epsilon|^2}{1+2\,\mathrm{Re}(\epsilon)+|\epsilon|^2}}$$
$$= \sqrt{\frac{4|\epsilon|^2}{1+3|\epsilon|^2}} = \left(\frac{3}{4}+\frac{1}{4|\epsilon|^2}\right)^{-\frac{1}{2}}, \tag{5.1.45}$$

where $\mathrm{Re}(\epsilon) = |\epsilon|^2$ is recalled from (5.1.39). We insert the expression for $|\epsilon|^2$ from above and obtain

$$\text{prob(no change after time } t) = \left(1+\left(\frac{t\,\delta P}{2M\,\delta X}\right)^2\right)^{-\frac{1}{2}}. \tag{5.1.46}$$

5.1.3 *Long-time and short-time behavior*

Two details are remarkable, namely what happens for very short times and what for very long times, meaning

$$\text{very short times:}\quad \left(\frac{t\,\delta P}{2M\,\delta X}\right) \ll 1\,,$$
$$\text{very long times:}\quad \left(\frac{t\,\delta P}{2M\,\delta X}\right) \ll 1\,. \tag{5.1.47}$$

For very long times, we note that "decay" of this kind is not exponential, rather the probability of having not yet decayed,

$$\text{prob(no change after long time t)} \cong \frac{2M\,\delta X}{t\,\delta P} \propto \frac{1}{t}\,, \tag{5.1.48}$$

decreases with a power law, here t^{-1}. This is a generic feature in the sense that such long-time behavior cannot be exponential; it must always be slower, but the precise dependence on t varies from circumstance to circumstance. Roughly speaking, one gets $\propto t^{-N}$, where N is the number

of degrees of freedom that are involved. All of this is a consequence of the fact that all physical Hamilton operators are bounded from below — they have a state of lowest energy, a ground state.

The short-time behavior,

$$\begin{aligned}\text{prob(no change after a short time t)} &\cong 1 - \frac{1}{2}\left(\frac{t\,\delta P}{2M\,\delta X}\right)^2 \\ &= 1 - \underbrace{(\cdots)}_{>0} t^2\,, \end{aligned} \tag{5.1.49}$$

is also generic and very much so because we always have a deviation that is proportional to t^2. The positive coefficient that sets the time scale has a particular physical significance that we shall explore now.

Let us recall what we calculated. The probability (5.1.43) in question results from the probability amplitude

$$\begin{aligned}\int \mathrm{d}x\,\psi_0(x)^*\psi(x,t) &= \int \mathrm{d}x\,\langle\ |x,0\rangle\langle x,t|\ \rangle \\ &= \langle\ |\left(\int \mathrm{d}x\,|x,0\rangle\langle x,t|\right)|\ \rangle = \langle\ |U_{0,t}|\ \rangle\,, \end{aligned} \tag{5.1.50}$$

which is the expectation value of the unitary evolution operator

$$U_{0,t} = \int \mathrm{d}x\,|x,0\rangle\langle x,t| \tag{5.1.51}$$

that links bras at one time to those at another time,

$$\langle x,0|U_{0,t} = \langle x,t|\,. \tag{5.1.52}$$

The Schrödinger equation obeyed by $\langle x,t|$,

$$\mathrm{i}\hbar\frac{\partial}{\partial t}\langle x,t| = \langle x,t|H(t)\,, \tag{5.1.53}$$

implies immediately that

$$U_{0,t} = \mathrm{e}^{-\mathrm{i}Ht/\hbar}\,, \tag{5.1.54}$$

provided that H does not depend on time,

$$\frac{\mathrm{d}}{\mathrm{d}t}H = \frac{\partial H}{\partial t} = 0\,. \tag{5.1.55}$$

The complications that arise for Hamilton operators with a parametric time dependence are of a technical nature, and we do not want to be distracted by them.

Accordingly, we have

$$\text{prob(no change after time } t) = \left| \left\langle \mathrm{e}^{-\mathrm{i}Ht/\hbar} \right\rangle \right|^2 , \tag{5.1.56}$$

which we now evaluate to second order in t. First, note that

$$\begin{aligned} \left\langle \mathrm{e}^{-\mathrm{i}Ht/\hbar} \right\rangle &= \left\langle \left(1 - \frac{\mathrm{i}}{\hbar} Ht + \frac{1}{2} \left(-\frac{\mathrm{i}}{\hbar} Ht \right)^2 + \cdots \right) \right\rangle \\ &= 1 - \frac{\mathrm{i}}{\hbar} \langle H \rangle \, t - \frac{1}{2} \frac{1}{\hbar^2} \left\langle H^2 \right\rangle t^2 + \cdots , \end{aligned} \tag{5.1.57}$$

and then that

$$\begin{aligned} \left| \left\langle \mathrm{e}^{-\mathrm{i}Ht/\hbar} \right\rangle \right|^2 &= \left(1 - \frac{1}{2} \frac{1}{\hbar^2} \left\langle H^2 \right\rangle t^2 \right)^2 + \left(\frac{1}{\hbar} \langle H \rangle \, t \right)^2 + \cdots \\ &= 1 - \frac{1}{\hbar^2} \Big(\left\langle H^2 \right\rangle - \langle H \rangle^2 \Big) t^2 + \cdots . \end{aligned} \tag{5.1.58}$$

We meet here the *energy spread*

$$\delta E = \delta H = \sqrt{\left\langle H^2 \right\rangle - \langle H \rangle^2} \tag{5.1.59}$$

and so arrive at

$$\begin{aligned} \text{prob(no change for small } t) &= 1 - \left(\frac{\delta E}{\hbar} t \right)^2 + O\big(t^4\big) \qquad (5.1.60) \\ &= 1 - \left(\frac{t}{T} \right)^2 + O\big(t^4\big) \quad \text{with} \quad T = \frac{\hbar}{\delta E} . \end{aligned}$$

The time scale for the initial "decay" is set by the energy spread δE; a large energy spread leads to rapid decay and a small one to slow decay.

No decay at all,

$$\text{prob(no change after } t) = 1 , \tag{5.1.61}$$

requires $\delta E = 0$. Then, the t^2 term disappears and also all higher-order terms. This is so because $\delta E = 0$ implies that $|\,\rangle$ is an eigenstate of the Hamilton operator,

$$H \big|\,\big\rangle = \big|\,\big\rangle E , \tag{5.1.62}$$

so that the outcome of an energy measurement can be predicted to be this eigenvalue E with certainty. Then,

$$\left\langle \mathrm{e}^{-\mathrm{i}Ht/\hbar} \right\rangle = \mathrm{e}^{-\mathrm{i}Et/\hbar} \tag{5.1.63}$$

and

$$\text{prob(no change after time } t) = \left| \mathrm{e}^{-\mathrm{i}Et/\hbar} \right|^2 = 1\,, \tag{5.1.64}$$

indeed there is no change at all.

Although it is quite obvious that, for any hermitian operator A, a vanishing spread, $\delta A = 0$, implies that the ket in question is an eigenket of A, let us give another argument for this fact, an argument of a somewhat geometrical nature. Given $A = A^\dagger$ and ket $|\,\rangle$ of unit length, $\langle\,|\,\rangle = 1$, then $A|\,\rangle$ is a new ket that, quite generally, has a component parallel to $|\,\rangle$ and one orthogonal to $|\,\rangle$,

$$A\big|\,\big\rangle = \big|\,\big\rangle a + \big|_\perp\big\rangle b \tag{5.1.65}$$

with coefficients a, b and $\langle\,|_\perp\rangle = 0$, $\langle_\perp|_\perp\rangle = 1$. The significance of a is immediate,

$$a = \Big\langle\,\Big|\Big(A\big|\,\big\rangle\Big) = \langle A\rangle\ , \tag{5.1.66}$$

it is the expectation value of A for ket $|\,\rangle$. It follows that a is real, whereas b can be complex.

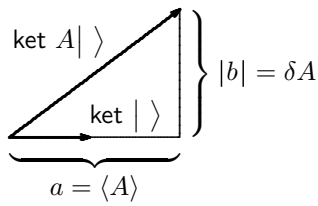
The significance of b is then revealed by considering the square of $A|\,\rangle$,

$$\Big(\langle\,|A\Big)\Big(A\big|\,\big\rangle\Big) = \Big\langle A^2 \Big\rangle = a^2 + b^* b\,, \tag{5.1.67}$$

so that $|b|^2$ is the variance of A,

$$|b|^2 = \Big\langle A^2 \Big\rangle - a^2 = \Big\langle A^2 \Big\rangle - \langle A\rangle^2 = (\delta A)^2\,.$$

We thus have this picture



(5.1.68)

and it follows immediately that $\delta A = 0$ implies

$$A\big|\,\big\rangle = \big|\,\big\rangle a\,, \tag{5.1.69}$$

that is, $|\,\rangle$ is an eigenket of A, and then $\left\langle \mathrm{e}^{\mathrm{i}\beta A}\right\rangle = \mathrm{e}^{\mathrm{i}\beta a}$ is a pure phase factor, $\left|\left\langle \mathrm{e}^{\mathrm{i}\beta A}\right\rangle\right| = 1$, for all real β. In the context of the above considerations of prob(no decay), it follows that there is no evolution whatsoever unless $\delta E \neq 0$. If the system is in an eigenstate of the Hamilton operator, $\delta E = 0$, then it does not evolve in time at all.

Comparing now the general result for prob(no change) in (5.1.60) with the particular example of (5.1.49),

$$1 - \frac{1}{2}\left(\frac{t\,\delta P}{2M\,\delta X}\right)^2 + O(t^4) \quad \text{with} \quad \delta X\,\delta P = \frac{1}{2}\hbar\,, \tag{5.1.70}$$

we infer that here

$$\frac{1}{\hbar}\delta E = \frac{1}{\sqrt{8}}\frac{\delta P}{M\,\delta X} \quad \text{or} \quad \delta E = \frac{1}{\sqrt{2}}\frac{(\delta P)^2}{M}\,. \tag{5.1.71}$$

As a check of consistency, we verify that this is indeed the energy spread for $H = P^2/(2M)$ and the initial wave function $\psi_0(x)$ in (5.1.28). Since

$$\delta E = \sqrt{\left\langle H^2\right\rangle - \langle H\rangle^{\,2}} = \frac{1}{2M}\sqrt{\left\langle P^4\right\rangle - \left\langle P^2\right\rangle^{\,2}} \tag{5.1.72}$$

involves expectation values of powers of the momentum operator, it will be convenient to employ the momentum wave function

$$\psi_0(p) = \frac{(2\pi)^{-\frac{1}{4}}}{\sqrt{\delta P}}\,\mathrm{e}^{-\left(\frac{1}{2}p/\delta P\right)^2} \tag{5.1.73}$$

rather than the position wave function (5.1.28); see (4.8.15). We have, in general,

$$\left\langle \left(P^2\right)^n\right\rangle = \int \mathrm{d}p\,\left(p^2\right)^n \frac{1}{\sqrt{2\pi}}\frac{1}{\delta P}\,\mathrm{e}^{-\frac{1}{2}(p/\delta P)^2}\,, \tag{5.1.74}$$

which are gaussian-type integrals of the form

$$\begin{aligned}\int \mathrm{d}p\,\left(p^2\right)^n \mathrm{e}^{-\gamma p^2} &= \left(-\frac{\partial}{\partial\gamma}\right)^n \int \mathrm{d}p\;\mathrm{e}^{-\gamma p^2} \\ = \left(-\frac{\partial}{\partial\gamma}\right)^n \sqrt{\frac{\pi}{\gamma}} &= \frac{1}{2}\frac{3}{2}\cdots\frac{2n-1}{2}\sqrt{\frac{\pi}{\gamma^{2n+1}}} = \frac{(2n)!}{4^n\,n!}\sqrt{\frac{\pi}{\gamma^{2n+1}}}\,.\end{aligned} \tag{5.1.75}$$

We need the $n = 1$ case in

$$\left\langle P^2 \right\rangle = \frac{1}{\sqrt{2\pi}} \frac{1}{\delta P} \frac{1}{2} \sqrt{\frac{\pi}{\left(\frac{1}{2}/(\delta P)^2\right)^3}} = (\delta P)^2 \tag{5.1.76}$$

(which could have been anticipated) and the $n = 2$ case in

$$\left\langle P^4 \right\rangle = \frac{1}{\sqrt{2\pi}} \frac{1}{\delta P} \frac{3}{4} \sqrt{\pi \big(2(\delta P)^2\big)^5} = 3(\delta P)^4 \,. \tag{5.1.77}$$

Together,

$$\delta\big(P^2\big) = \sqrt{\left\langle P^4 \right\rangle - \left\langle P^2 \right\rangle^2} = \sqrt{3(\delta P)^4 - \big((\delta P)^2\big)^2} = \sqrt{2}\,(\delta P)^2 \,, \tag{5.1.78}$$

and

$$\delta E = \frac{1}{2M} \sqrt{2} (\delta P)^2 = \frac{1}{\sqrt{2}} \frac{(\delta P)^2}{M} \tag{5.1.79}$$

follows, indeed.

A final remark on these matters is the observation that we have, in a sense, seen the quadratic initial time dependence before, namely when discussing many intermediate Stern–Gerlach measurements in Section 2.6. The general statement that we can formulate now is about many intermediate control measurements that ask if some evolution has happened. If the total duration T is broken up into n intervals of duration $t = T/n$, then the probability of no change all together will be

$$\begin{aligned} &\text{prob(no change in } T\text{, with } n \text{ controls)} \\ &= \text{prob(no change in } t = T/n)^n \,. \end{aligned} \tag{5.1.80}$$

Now, if n is so large that the quadratic short-time approximation applies to $t = T/n$, then

$$\begin{aligned} \text{prob(no change in } T\text{, with } n \text{ controls)} &\cong \left(1 - \left(\frac{\delta E}{\hbar} \frac{T}{n}\right)^2\right)^n \\ \cong \left(\mathrm{e}^{-\left(\frac{\delta E}{\hbar} \frac{T}{n}\right)^2}\right)^n &= \mathrm{e}^{-\frac{1}{n}\left(\delta E \frac{T}{\hbar}\right)^2} \cong 1 - \frac{1}{n}\left(\delta E \frac{T}{\hbar}\right)^2 . \end{aligned} \tag{5.1.81}$$

This states an effective slowdown because the time constant for decay, $\sqrt{n}\,\hbar/\delta E$, is $\sqrt{n}$ times the original one. For n so large that $\frac{1}{n}\big(\delta E\, T/\hbar\big)^2$ is of the desired small size, the evolution is thus effectively halted by the many

interrupting control measurements. As mentioned in Section 2.6, this quantum Zeno effect is a real physical phenomenon that has been demonstrated repeatedly in a variety of experiments.

5.1.4 *Interlude: General position-dependent force*

In the presence of a force, $F = -\frac{\partial}{\partial x}V(x)$, derived from a potential energy $V(x)$, the Hamilton operator is of the typical form

$$H = \frac{1}{2M}P^2 + V(X)\,. \tag{5.1.82}$$

The Heisenberg equations of motion are then

$$\frac{\mathrm{d}}{\mathrm{d}t}X = \frac{1}{\mathrm{i}\hbar}[X,H] = \frac{1}{\mathrm{i}\hbar}\left[X,\frac{1}{2M}P^2\right] = \frac{1}{M}P \tag{5.1.83}$$

and

$$\frac{\mathrm{d}}{\mathrm{d}t}P = \frac{1}{\mathrm{i}\hbar}[P,H] = \frac{1}{\mathrm{i}\hbar}\big[P,V(X)\big] = -\frac{\partial V}{\partial X}(X) \tag{5.1.84}$$

or

$$\frac{\mathrm{d}}{\mathrm{d}t}P = F(X)\,. \tag{5.1.85}$$

The latter commutator is an example of

$$\big[f(X),P\big] = \mathrm{i}\hbar\frac{\partial f}{\partial X}(X)\,, \tag{5.1.86}$$

that is, we differentiate with respect to X by taking the commutator with P. This equation can be proven in various ways, simplest perhaps by applying the operators to x bras,

$$\begin{aligned}\big\langle x\big|\big[f(X),P\big] &= \big\langle x\big|\big(f(X)P - Pf(X)\big)\\ &= \Bigg(f(x)\frac{\hbar}{\mathrm{i}}\frac{\partial}{\partial x} - \frac{\hbar}{\mathrm{i}}\underbrace{\frac{\partial}{\partial x}f(x)}_{=f(x)\frac{\partial}{\partial x}+\frac{\partial f(x)}{\partial x}}\Bigg)\big\langle x\big|\\ &= \mathrm{i}\hbar\frac{\partial f(x)}{\partial x}\big\langle x\big| = \big\langle x\big|\mathrm{i}\hbar\frac{\partial f(X)}{\partial X}\,,\end{aligned} \tag{5.1.87}$$

and then the commutator identity follows from the completeness of the bras $\langle x|$.

The pair of Heisenberg equations of motion (5.1.83) and (5.1.84) constitute, as a rule, a coupled set of *nonlinear* differential equation for the

operators $X(t)$ and $P(t)$, because an arbitrary function $V(X)$ is involved, and we do not have methods for solving them explicitly. By "solving" we mean to express $X(t)$ and $P(t)$ in terms of $X(0)$ and $P(0)$, as we managed to do for the force-free particle case of $V = 0$. There are, however, various methods for obtaining approximate solutions in a systematic manner, part and parcel of *perturbation theory*, which is beyond the scope of *Basic Matters* but is discussed in *Simple Systems* and *Perturbed Evolution*.

One can always retreat to the formal solutions

$$\begin{aligned} X(t) &= \mathrm{e}^{\mathrm{i}Ht/\hbar} X(0)\, \mathrm{e}^{-\mathrm{i}Ht/\hbar}\,, \\ P(t) &= \mathrm{e}^{\mathrm{i}Ht/\hbar} P(0)\, \mathrm{e}^{-\mathrm{i}Ht/\hbar}\,, \end{aligned} \tag{5.1.88}$$

which solve the Heisenberg equations of motion, provided that H itself does not depend on time

$$\frac{\mathrm{d}H}{\mathrm{d}t} = 0\,. \tag{5.1.89}$$

As we know, this is the requirement that there is no parametric time dependence in H, $\partial H/\partial t = 0$. More generally, any operator function $F\big(X(t), P(t)\big)$ can be related to its $t = 0$ version by

$$F\big(X(t), P(t)\big) = \mathrm{e}^{\mathrm{i}Ht/\hbar} F\big(X(0), P(0)\big)\, \mathrm{e}^{-\mathrm{i}Ht/\hbar}\,, \tag{5.1.90}$$

and the unitary evolution operator $\mathrm{e}^{-\mathrm{i}Ht/\hbar}$ also provides the time-dependent bras,

$$\big\langle \ldots, t\big| = \big\langle \ldots, 0\big|\, \mathrm{e}^{-\mathrm{i}Ht/\hbar}\,, \tag{5.1.91}$$

as we have already observed in (5.1.52)–(5.1.54).

All of this directs our attention to (3.8.4), that is,

$$\mathrm{e}^{-\mathrm{i}Ht/\hbar} = \sum_j \big|E_j\big\rangle\, \mathrm{e}^{-\mathrm{i}E_j t/\hbar} \big\langle E_j\big|\,, \tag{5.1.92}$$

where the sum is over the eigenvalues of H,

$$H\big|E_j\big\rangle = \big|E_j\big\rangle E_j\,, \qquad \big\langle E_j\big| H = E_j \big\langle E_j\big|\,, \tag{5.1.93}$$

physically speaking the *eigenenergies* of the system. This overwhelming importance for time evolution gives a very special role to the eigenvalues and eigenstates of the Hamilton operator and, therefore, a variety of techniques have been developed for determining the eigenenergies and the energy eigenstates with high precision. Some of these methods are discussed in *Simple*

Systems. They are, of course, also applicable to other hermitian operators, but most of the time, it is the Hamilton operator that is both of special interest and particularly difficult to deal with.

5.1.5 *Energy eigenstates*

For the force-free motion, we have $H = P^2/(2M)$ and the time-independent Schrödinger equation (that is, the eigenvalue equation for H) reads

$$\frac{1}{2M}P^2\big|E\big\rangle = \big|E\big\rangle E\,. \tag{5.1.94}$$

Clearly, the eigenkets of the momentum operator P are also eigenkets of its square $P^2 = 2MH$ and, therefore, we may identify one with the other,

$$\big|E\big\rangle \propto \big|p\big\rangle \quad \text{with} \quad E = \frac{1}{2M}p^2\,. \tag{5.1.95}$$

We encounter here, in a simple context, a quite ubiquitous detail, namely that there is more than one state for a given energy. It is rather common that energy eigenvalues are *degenerate* by which we mean just that: there is more than one state for the given energy eigenvalue. Such degeneracies are always the consequence of some symmetry property. In the case of the force-free motion, it is the *reflection symmetry*: it does not matter whether the atom moves to the left ($p < 0$) or to the right ($p > 0$).

We thus have two eigenstates for each energy eigenvalue $E > 0$,

$$\big|E,+\big\rangle \propto \big|p = +\sqrt{2ME}\,\big\rangle \tag{5.1.96}$$

and

$$\big|E,-\big\rangle \propto \big|p = -\sqrt{2ME}\,\big\rangle\,, \tag{5.1.97}$$

and a single eigenstate for $E = 0$,

$$\big|E=0\big\rangle \propto \big|p=0\big\rangle\,. \tag{5.1.98}$$

Since all p kets appear here, which constitute a complete set, the kets $|E,\alpha\rangle$ with $E > 0$ and $\alpha = \pm$ together with the ket for $E = 0$ are complete, too. In fact, it doesn't matter if we omit the $E = 0$ state; the remaining ones are still complete because in integrals such as

$$\big\langle 1\big|2\big\rangle = \int_{-\infty}^{\infty} \mathrm{d}p\; \psi_1(p)^*\psi_2(p)\,, \tag{5.1.99}$$

the omission of the point contribution at $p = 0$ does not affect the value. We shall, therefore, simplify matters by ignoring the $E = 0$ ket.

All "+" kets are automatically orthogonal to the "−" kets,

$$\big\langle E, + \big| E', - \big\rangle = 0\,, \tag{5.1.100}$$

and since E is a continuous label (all $E > 0$ are permissible), we have to normalize the energy eigenkets to a delta function

$$\begin{aligned} \big\langle E, + \big| E', + \big\rangle &= \delta(E - E')\,, \\ \big\langle E, - \big| E', - \big\rangle &= \delta(E - E')\,. \end{aligned} \tag{5.1.101}$$

With this normalization, the identity is decomposed as

$$1 = \sum_{\alpha=\pm} \int_0^\infty \mathrm{d}E\, \big| E, \alpha \big\rangle \big\langle E, \alpha \big|\,. \tag{5.1.102}$$

We verify this completeness relation by applying it to $|E', \beta\rangle$ with $E' > 0$ and $\beta = \pm$,

$$\int_0^\infty \mathrm{d}E \sum_\alpha \big| E, \alpha \big\rangle \underbrace{\big\langle E, \alpha \big| E', \beta \big\rangle}_{= \delta_{\alpha\beta}\delta(E - E')} = \big| E', \beta \big\rangle\,. \tag{5.1.103}$$

Upon writing

$$\big| E, \pm \big\rangle = \big| p = \pm\sqrt{2ME}\,\big\rangle a_\pm(E) \tag{5.1.104}$$

with the normalizing factor $a_\pm(E)$ to be determined, we have

$$\begin{aligned} \big\langle E, \alpha \big| E', \beta \big\rangle &= a_\alpha(E)^* \big\langle p = \alpha\sqrt{2ME} \big| p = \beta\sqrt{2ME'}\,\big\rangle a_\beta(E') \\ &= a_\alpha(E)^* \underbrace{\delta\Big(\alpha\sqrt{2ME} - \beta\sqrt{2ME'}\Big)}_{= 0 \text{ if } \alpha \neq \beta} a_\beta(E') \\ &= a_\alpha(E)^* \delta_{\alpha\beta}\, \delta\Big(\sqrt{2ME} - \sqrt{2ME'}\Big) a_\beta(E')\,. \end{aligned} \tag{5.1.105}$$

The delta function appearing here has a vanishing argument only when $E = E'$, so it must be expressible in terms of $\delta(E - E')$. Let us, therefore, introduce $\mathcal{E} = E - E'$ for the difference,

$$\delta\Big(\sqrt{2ME} - \sqrt{2ME'}\Big) = \delta\Big(\sqrt{2M(E' + \mathcal{E})} - \sqrt{2ME'}\Big) \tag{5.1.106}$$

and note that only the immediate vicinity of $\mathcal{E} = 0$ is relevant. But there we have

$$\sqrt{2M(E' + \mathcal{E})} = \sqrt{2ME'} + \sqrt{\frac{M}{2E'}}\,\mathcal{E} \tag{5.1.107}$$

so that

$$\begin{aligned}\delta\Big(\sqrt{2ME} - \sqrt{2ME'}\Big) = \delta\Big(\sqrt{M/(2E')}\,\mathcal{E}\Big) &= \sqrt{2E'/M}\,\delta(\mathcal{E}) \\ &= \sqrt{2E'/M}\,\delta(E - E')\,. \end{aligned} \tag{5.1.108}$$

The second equality here is an illustration of

$$\delta(\lambda x) = \frac{1}{\lambda}\delta(x) \quad \text{for} \quad \lambda > 0\,, \tag{5.1.109}$$

an identity that is used very often. Itself, this identity is a special case of

$$\delta\big(f(x)\big) = \sum_j \big|f'(x_j)\big|^{-1}\delta(x - x_j)\,, \tag{5.1.110}$$

where $f(x)$ is a real-valued function of x that has simple zeros at the x_js and $f'(x)$ denotes the derivative of $f(x)$; see Exercise 103.

So, we have

$$\begin{aligned}\delta_{\alpha\beta}\,\delta(E - E') &= \big\langle E, \alpha\big|E', \beta\big\rangle \\ &= a_\alpha(E)^*\delta_{\alpha\beta}\sqrt{2E'/M}\,\delta(E - E')a_\beta(E') \end{aligned} \tag{5.1.111}$$

and conclude that

$$\big|a_\alpha(E)\big|^2\sqrt{2E/M} = 1 \tag{5.1.112}$$

is needed. Since there is no point in introducing arbitrary phase factors, we just take $a_\alpha(E) = (2E/M)^{-\frac{1}{4}}$ and get

$$\big|E, \alpha\big\rangle = \big|p = \alpha\sqrt{2ME}\,\big\rangle(2E/M)^{-\frac{1}{4}} \tag{5.1.113}$$

for the normalized eigenkets of the force-free Hamilton operator.

Since the energies are degenerate and we have here two states to each energy, we can take any pairwise orthogonal superpositions of same-energy eigenstates and get other sets of eigenstates, sets that are equally good. For example, the kets

$$\begin{aligned}\big|E, \text{even}\big\rangle &= \Big(\big|E, +\big\rangle + \big|E, -\big\rangle\Big)\Big/\sqrt{2}\,, \\ \big|E, \text{odd}\big\rangle &= \Big(\big|E, +\big\rangle - \big|E, -\big\rangle\Big)\Big/\big(\mathrm{i}\sqrt{2}\,\big) \end{aligned} \tag{5.1.114}$$

will serve the purpose just as well, inasmuch as they are orthonormal

$$\big\langle E, a\big|E', b\big\rangle = \delta_{ab}\,\delta(E-E')\,; \quad a, b = \text{even, odd}, \tag{5.1.115}$$

and complete

$$\int_0^\infty \mathrm{d}E\,\Big(\big|E,\text{even}\big\rangle\big\langle E,\text{even}\big| + \big|E,\text{odd}\big\rangle\big\langle E,\text{odd}\big|\Big) = 1\,. \tag{5.1.116}$$

5.2 Constant force

5.2.1 *Energy eigenstates*

Now, getting to more complicated matters, let us consider motion under the influence of a constant force F so that

$$H = \frac{1}{2M}P^2 - FX\,, \quad F = \text{constant}, \tag{5.2.1}$$

is the Hamilton operator. Indeed, Heisenberg's equations of motion,

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}P &= \frac{1}{\mathrm{i}\hbar}[P,H] = F\,,\\ \frac{\mathrm{d}}{\mathrm{d}t}X &= \frac{1}{\mathrm{i}\hbar}[X,H] = \frac{1}{M}P\,,\end{aligned} \tag{5.2.2}$$

are those of a particle of mass M exposed to a force F. The energy eigenvalues are determined by

$$H\big|E\big\rangle = \big|E\big\rangle E\,. \tag{5.2.3}$$

With $P \to \dfrac{\hbar}{\mathrm{i}}\dfrac{\partial}{\partial x}$ and $X \to x$, this is

$$\left(-\frac{\hbar^2}{2M}\frac{\partial^2}{\partial x^2} - Fx\right)\psi_E(x) = E\psi_E(x) \tag{5.2.4}$$

for the position wave function $\psi_E(x) = \langle x|E\rangle$, and with $P \to p$, $X \to \mathrm{i}\hbar\dfrac{\partial}{\partial p}$, we get

$$\left(\frac{p^2}{2M} - \mathrm{i}\hbar F\frac{\partial}{\partial p}\right)\psi_E(p) = E\psi_E(p) \tag{5.2.5}$$

for the momentum wave function $\psi_E(p) = \langle p|E\rangle$.

We have here a situation in which it is easier to solve the momentum variant of the time-independent Schrödinger equation. Upon introducing an integrating factor,

$$\frac{p^2}{2M} - \mathrm{i}\hbar F\frac{\partial}{\partial p} = \mathrm{e}^{-\frac{\mathrm{i}}{\hbar}\frac{p^3}{6MF}}\left(-\mathrm{i}\hbar F\frac{\partial}{\partial p}\right)\mathrm{e}^{\frac{\mathrm{i}}{\hbar}\frac{p^3}{6MF}}\,, \tag{5.2.6}$$

we have

$$\frac{\partial}{\partial p}\Big(\,\mathrm{e}^{\frac{\mathrm{i}}{\hbar}\frac{p^3}{6MF}}\psi_E(p)\Big) = \frac{\mathrm{i}E}{\hbar F}\Big(\,\mathrm{e}^{\frac{\mathrm{i}}{\hbar}\frac{p^3}{6MF}}\psi_E(p)\Big) \tag{5.2.7}$$

so that

$$\mathrm{e}^{\frac{\mathrm{i}}{\hbar}\frac{p^3}{6MF}}\psi_E(p) = a(E)\,\mathrm{e}^{\frac{\mathrm{i}}{\hbar}\frac{Ep}{F}} \tag{5.2.8}$$

or

$$\psi_E(p) = a(E)\,\mathrm{e}^{-\frac{\mathrm{i}}{\hbar}\frac{p^3}{6MF}}\,\mathrm{e}^{\frac{\mathrm{i}}{\hbar}\frac{Ep}{F}}\,. \tag{5.2.9}$$

The p-independent factor $a(E)$ is fixed, to within an irrelevant phase factor, by the normalization condition

$$\big\langle E\big|E'\big\rangle = \int \mathrm{d}p\;\psi_E(p)^*\psi_{E'}(p) = \delta(E-E')\,. \tag{5.2.10}$$

There is, by the way, no restriction on the possible values for E; all real values are eigenvalues of H. Clearly, we are dealing with an overidealized, oversimplified model. Real physical Hamilton operators are always bounded from below. The idealization is that of a truly constant force, constant over the whole range of x, not just a limited, perhaps large but limited, range.

The orthonormality requirement is more explicitly

$$\begin{aligned}\delta(E-E') &= \int \mathrm{d}p\,a(E)^*\,\mathrm{e}^{\frac{\mathrm{i}}{\hbar}\frac{p^3}{6MF}}\,\mathrm{e}^{-\frac{\mathrm{i}}{\hbar}\frac{Ep}{F}}\,a(E')\,\mathrm{e}^{-\frac{\mathrm{i}}{\hbar}\frac{p^3}{6MF}}\,\mathrm{e}^{\frac{\mathrm{i}}{\hbar}\frac{E'p}{F}}\\ &= a(E)^*a(E')\int \mathrm{d}p\;\mathrm{e}^{\frac{\mathrm{i}}{\hbar}\frac{E'-E}{F}p}\\ &= a(E)^*a(E')\,2\pi\hbar\,\delta\big((E'-E)/F\big)\\ &= 2\pi\hbar\,\big|F\big|\,\big|a(E)\big|^2\,\delta(E'-E)\,,\end{aligned} \tag{5.2.11}$$

where we recall the Fourier representation of the delta function in (4.6.14) and the identity (5.1.109). Opting for the simplest phase convention, $a(E) > 0$, we thus have

$$a(E) = \frac{1}{\sqrt{2\pi\hbar\,\big|F\big|}} \tag{5.2.12}$$

and arrive at

$$\psi_E(p) = \left(2\pi\hbar\,\big|F\big|\right)^{-\frac{1}{2}} \mathrm{e}^{-\frac{\mathrm{i}}{\hbar}\frac{p^3}{6MF}}\,\mathrm{e}^{\frac{\mathrm{i}}{\hbar}\frac{Ep}{F}}\,. \tag{5.2.13}$$

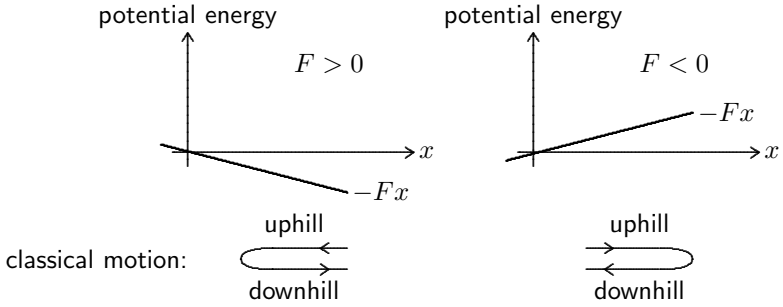
The position wave function is then available by Fourier transformation,

$$\begin{aligned}\psi_E(x) = \big\langle x\big|E\big\rangle &= \int \mathrm{d}p\,\big\langle x\big|p\big\rangle\big\langle p\big|E\big\rangle = \int \mathrm{d}p\,\frac{\mathrm{e}^{\mathrm{i}xp/\hbar}}{\sqrt{2\pi\hbar}}\psi_E(p)\\ &= \big|F\big|^{-\frac{1}{2}}\int\frac{\mathrm{d}p}{2\pi\hbar}\,\mathrm{e}^{\mathrm{i}(x\,+\,E/F)p/\hbar}\,\mathrm{e}^{-\frac{\mathrm{i}}{\hbar}\frac{p^3}{6MF}}\,,\end{aligned} \tag{5.2.14}$$

and the integration over p yields $\psi_E(x)$ in terms of the Airy* function; see Exercise 106.

5.2.2 *Limit of no force*

We shall not explore the properties of $\psi_E(x)$ in detail but wish to see how the force-free wave functions of Section 5.1.5 emerge in the limit $F \to 0$. This limit is, in fact, a bit delicate because the physical situations of $F > 0$ and $F < 0$ are quite different. In the first case, all solutions will refer to particles moving downhill to the right, whereas $F < 0$ has downhill motion to the left with a clear discontinuity at $F = 0$:



(5.2.15)

Indeed, $\psi_E(p)$ is singular for $F = 0$, with a singularity of the delta function kind. Therefore, it will be technically a bit easier to study the $F \to 0$ limit at the position wave function

$$\psi_E(x) = \frac{1}{\sqrt{|F|}}\int\frac{\mathrm{d}p}{2\pi\hbar}\,\mathrm{e}^{\mathrm{i}W(p)/\hbar} \tag{5.2.16}$$

*George Bidell Airy (1801–1892)

with the p-dependent action variable

$$W(p) = \left(x + \frac{E}{F}\right)p - \frac{p^3}{6MF}\,. \tag{5.2.17}$$

The phase $W(p)/\hbar$ of the exponential function changes very rapidly as a function of p when $F \cong 0$, except at the *points of stationary phase* where $\dfrac{\partial W}{\partial p} = 0$. Close to these points the otherwise rapidly oscillating $\mathrm{e}^{\mathrm{i}W/\hbar}$ factor is actually constant so that p values in the vicinity contribute in phase. In other words, there is constructive interference near points of stationary phase and destructive interference anywhere else.

The points of stationary phase are the solutions of the equation

$$\frac{\partial W}{\partial p}(p) = 0 \quad \text{or} \quad \frac{p^2}{2MF} = x + \frac{E}{F}\,, \tag{5.2.18}$$

that is,

$$p = \pm\overline{p}\,, \quad \overline{p} = \sqrt{2ME + 2MFx}\,. \tag{5.2.19}$$

Near the p values of stationary phase we have

$$W(p) = W\Big(\pm\overline{p} + (p \mp \overline{p})\Big) \cong W(\pm\overline{p}) + \frac{1}{2}\frac{\partial^2 W}{\partial p^2}(\pm\overline{p})(p \mp \overline{p})^2\,, \tag{5.2.20}$$

terminating the Taylor expansion about $p = \pm\overline{p}$ at the quadratic term; then,

$$W(\pm\overline{p}) = \pm\left(x + \frac{E}{F} - \frac{\overline{p}^2}{6MF}\right)\overline{p} = \pm\frac{\overline{p}^3}{3MF} = \pm W(\overline{p}) \tag{5.2.21}$$

and

$$\frac{\partial^2 W}{\partial p^2}(\pm\overline{p}) = \mp\frac{\overline{p}}{MF}\,. \tag{5.2.22}$$

In the p integral for $\psi_E(x)$, we thus have two major contributions for $p \cong +\overline{p}$ and $p \cong -\overline{p}$, which we wish to pick out for small F values, eventually taking the limit $F \to 0$. But then we are restricted to $E > 0$ because otherwise

$$\overline{p} \xrightarrow[F\,\to\,0]{} \sqrt{2ME} \tag{5.2.23}$$

turns imaginary and $\mathrm{e}^{\mathrm{i}W(p)/\hbar}$ becomes an exponentially decreasing quantity. This mathematical detail reminds us that we have only $E > 0$ for force-free motion ($F = 0$) but arbitrary E for $F \neq 0$, however small F is.

Now, for $E>0$, we have contributions for $p\cong\overline{p}$ and $p\cong-\overline{p}$. It is convenient to parameterize the integrals as $p=\overline{p}+q$ and $p=-\overline{p}+q$, respectively, so that

$$\begin{aligned}\psi_E(x) &\cong \frac{1}{\sqrt{|F|}}\int\frac{\mathrm{d}q}{2\pi\hbar}\Big(\mathrm{e}^{\frac{\mathrm{i}}{\hbar}(W(\overline{p})-\frac{\overline{p}}{2MF}q^2)}+\mathrm{e}^{-\frac{\mathrm{i}}{\hbar}(W(\overline{p})-\frac{\overline{p}}{2MF}q^2)}\Big)\\ &=\frac{1}{\sqrt{|F|}}\frac{1}{2\pi\hbar}\left(\mathrm{e}^{\frac{\mathrm{i}}{\hbar}W(\overline{p})}\sqrt{\frac{\pi}{+\mathrm{i}\frac{\overline{p}}{2\hbar MF}}}+\mathrm{e}^{-\frac{\mathrm{i}}{\hbar}W(\overline{p})}\sqrt{\frac{\pi}{-\mathrm{i}\frac{\overline{p}}{2\hbar MF}}}\right)\\ &=\frac{1}{\sqrt{2\pi\hbar}}\left(\sqrt{\mathrm{sgn}(F)\frac{M}{\mathrm{i}\overline{p}}}\,\mathrm{e}^{\frac{\mathrm{i}}{\hbar}W(\overline{p})}+\sqrt{\mathrm{sgn}(F)\frac{\mathrm{i}M}{\overline{p}}}\,\mathrm{e}^{-\frac{\mathrm{i}}{\hbar}W(\overline{p})}\right). \quad (5.2.24)\end{aligned}$$

Here, $\mathrm{sgn}(F)=F/|F|$ is the sign of F and

$$\begin{aligned}W(\overline{p})=\frac{\overline{p}^3}{3MF}&=\frac{1}{3MF}\sqrt{2ME+2MFx}^3\\ &=\frac{1}{3MF}\sqrt{2ME}^3\Big(1+\frac{3}{2}\frac{Fx}{E}+O\big(F^2\big)\Big)\\ &\cong\frac{(2ME)^{\frac{3}{2}}}{3MF}+\sqrt{2ME}\,x\,, \quad (5.2.25)\end{aligned}$$

when keeping only the two leading terms for small F. We also have

$$\overline{p}/M\to\sqrt{2E/M}\quad\text{for}\quad F\to0 \quad (5.2.26)$$

so that

$$\begin{aligned}\psi_E(x)\cong\frac{1}{\sqrt{2\pi\hbar}}(2E/M)^{-\frac{1}{4}}\Bigg(&\sqrt{\frac{1}{\mathrm{i}}\,\mathrm{sgn}(F)}\,\mathrm{e}^{\mathrm{i}\phi}\,\mathrm{e}^{\frac{\mathrm{i}}{\hbar}\sqrt{2ME}\,x}\\ &+\sqrt{\mathrm{i}\,\mathrm{sgn}(F)}\,\mathrm{e}^{-\mathrm{i}\phi}\,\mathrm{e}^{-\frac{\mathrm{i}}{\hbar}\sqrt{2ME}\,x}\Bigg) \quad (5.2.27)\end{aligned}$$

with $\phi=(3MF\hbar)^{-\frac{1}{2}}(2ME)^{\frac{3}{2}}$, which is an x-independent phase factor.

Noting that

$$\begin{aligned}\frac{1}{\sqrt{2\pi\hbar}}\,\mathrm{e}^{\pm\mathrm{i}\sqrt{2ME}x/\hbar}&=\big\langle x\big|p=\pm\sqrt{2ME}\,\big\rangle\\ &=\big\langle x\big|E,\pm\big\rangle(2E/M)^{\frac{1}{4}} \quad (5.2.28)\end{aligned}$$

with $|E,\alpha\rangle$ from (5.1.113), we thus have

$$\underbrace{|E\rangle}_{\text{force } F} \cong \underbrace{|E,+\rangle}_{F=0} \mathrm{e}^{\mathrm{i}\phi}\sqrt{\frac{1}{\mathrm{i}}\operatorname{sgn}(F)} + \underbrace{|E,-\rangle}_{F=0} \mathrm{e}^{-\mathrm{i}\phi}\sqrt{\mathrm{i}\operatorname{sgn}(F)} \tag{5.2.29}$$

as the $F \cong 0$ approximation. We recognize the $F = 0$ energy eigenstates, multiplied by F-dependent, largely irrelevant phase factors.

5.3 Harmonic oscillator

5.3.1 *Energy eigenstates: Power-series method*

After the force-free motion, $V(X) = 0$, the motion guided by a constant force, $V(X) = -FX$, we now come to the *harmonic oscillator*, $V(X) = \frac{1}{2}M\omega^2X^2$, for which the Hamilton operator is

$$H = \frac{1}{2M}P^2 + \frac{1}{2}M\omega^2X^2\,. \tag{5.3.1}$$

The Heisenberg equations of motion,

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}P &= \frac{1}{\mathrm{i}\hbar}[P,H] = -\frac{\partial}{\partial X}H = -M\omega^2X\,,\\ \frac{\mathrm{d}}{\mathrm{d}t}X &= \frac{1}{\mathrm{i}\hbar}[X,H] = \frac{\partial}{\partial P}H = \frac{1}{M}P\,,\end{aligned} \tag{5.3.2}$$

can be iterated to produce

$$\left(\frac{\mathrm{d}}{\mathrm{d}t}\right)^2 X = \frac{1}{M}\frac{\mathrm{d}}{\mathrm{d}t}P = -\omega^2X\,, \tag{5.3.3}$$

which is the second-order equation of motion — Newton's "acceleration = force/mass" — of a harmonic oscillator, indeed.

The time-independent Schrödinger equation for position wave functions $\psi_E(x) = \langle x|E\rangle$ reads

$$\left(-\frac{\hbar^2}{2M}\frac{\partial^2}{\partial x^2} + \frac{1}{2}M\omega^2x^2\right)\psi_E(x) = E\psi_E(x)\,. \tag{5.3.4}$$

Since the oscillator is characterized by its natural (circular) frequency ω, it is expedient to measure the energy as a multiple of $\hbar\omega$, and the parameterization

$$\frac{E}{\hbar\omega} = n + \frac{1}{2} \tag{5.3.5}$$

will turn out to be particularly useful. Thus, the equation for eigenvalue E of H will turn into an equation for eigenvalue n of $H/(\hbar\omega) - \frac{1}{2}$,

$$\left(-\frac{\hbar}{2M\omega}\frac{\partial^2}{\partial x^2} + \frac{1}{2}\frac{M\omega}{\hbar}x^2 - \frac{1}{2}\right)\psi_n = n\psi_n \tag{5.3.6}$$

with $\psi_n = \psi_E$ for $E = \hbar\omega\big(n + \frac{1}{2}\big)$. The combination of factors

$$\frac{M\omega}{\hbar}x^2 = q^2 \tag{5.3.7}$$

is dimensionless ($M\omega x$ has the metrical dimension of momentum) so that

$$q = \sqrt{\frac{M\omega}{\hbar}}\, x \tag{5.3.8}$$

is the distance x expressed in the natural length unit $\sqrt{\hbar/(M\omega)}$ of the oscillator. Then

$$\frac{\hbar}{M\omega}\frac{\partial^2}{\partial x^2} = \frac{\partial^2}{\partial q^2} \tag{5.3.9}$$

and we regard ψ_n as a function of q,

$$\psi_n(q) \propto \psi_E(x) \quad \text{for} \quad E = \hbar\omega\left(n + \tfrac{1}{2}\right) \quad \text{and} \quad x = \sqrt{\frac{\hbar}{M\omega}}\, q\,. \tag{5.3.10}$$

We do not write $\psi_n(q) = \psi_E(x)$ because we may find it convenient, at some later stage, to normalize $\psi_n(q)$ differently than $\psi_E(x)$.

At this stage, we have

$$\left(-\frac{\mathrm{d}^2}{\mathrm{d}q^2} + q^2 - (2n+1)\right)\psi_n(q) = 0\,. \tag{5.3.11}$$

Following a method that was systematized by Schrödinger, we first consider the range of very large q,

$$q \gg 1: \quad \frac{\mathrm{d}^2}{\mathrm{d}q^2}\psi_n(q) \cong q^2\psi_n(q)\,, \tag{5.3.12}$$

and note that

$$\psi_n(q) = \mathrm{e}^{\frac{1}{2}q^2} \quad \text{or} \quad \mathrm{e}^{-\frac{1}{2}q^2} \tag{5.3.13}$$

would solve these equations (for large q values). Eventually, we shall need to normalize the solution and, therefore, we must insist on a finite integral

of its square,

$$\int \mathrm{d}q \left|\psi_n(q)\right|^2 < \infty\,. \tag{5.3.14}$$

It follows that $\mathrm{e}^{+\frac{1}{2}q^2}$ is not an option, and we are led to the *ansatz*

$$\psi_n(q) = \chi_n(q)\,\mathrm{e}^{-\frac{1}{2}q^2}\,, \tag{5.3.15}$$

that is, we put aside the large-q exponential and regard $\chi_n(q)$ as the function to be determined.

Since

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}q}\psi_n &= \mathrm{e}^{-\frac{1}{2}q^2}\left(\frac{\mathrm{d}}{\mathrm{d}q} - q\right)\chi_n\,,\\
\left(\frac{\mathrm{d}}{\mathrm{d}q}\right)^2\psi_n &= \mathrm{e}^{-\frac{1}{2}q^2}\left(\frac{\mathrm{d}^2}{\mathrm{d}q^2} - 2q\frac{\mathrm{d}}{\mathrm{d}q} + q^2 - 1\right)\chi_n\\
&= \big(q^2 - (2n+1)\big)\psi_n\\
&= \mathrm{e}^{-\frac{1}{2}q^2}\big(q^2 - 2n - 1\big)\chi_n\,,
\end{aligned} \tag{5.3.16}$$

we get

$$\left(\frac{\mathrm{d}^2}{\mathrm{d}q^2} - 2q\frac{\mathrm{d}}{\mathrm{d}q} + 2n\right)\chi_n(q) = 0 \tag{5.3.17}$$

for the differential equation obeyed by $\chi_n(q)$. It contains, of course, the unknown eigenvalue $n = E/(\hbar\omega) - \frac{1}{2}$.

We write $\chi_n(q)$ as a power series in q,

$$\chi_n(q) = \sum_{k=0}^{\infty} \frac{1}{k!} a_k^{(n)} q^k, \tag{5.3.18}$$

so that

$$\frac{\mathrm{d}^2}{\mathrm{d}q^2}\chi_n(q) = \sum_{k=2}^{\infty} \frac{1}{(k-2)!} a_k^{(n)} q^{k-2} = \sum_{k=0}^{\infty} \frac{1}{k!} a_{k+2}^{(n)} q^k \tag{5.3.19}$$

and

$$q\frac{\mathrm{d}}{\mathrm{d}q}\chi_n(q) = \sum_{k=0}^{\infty} \frac{1}{k!} k a_k^{(n)} q^k\,. \tag{5.3.20}$$

Together they give us

$$a_{k+2}^{(n)} - 2k a_k^{(n)} + 2n a_k^{(n)} = 0 \tag{5.3.21}$$

for the coefficient of q^k in (5.3.17), which is a simple recurrence relation for the coefficients $a_k^{(n)}$,

$$a_{k+2}^{(n)} = 2(k-n)a_k^{(n)} \,. \tag{5.3.22}$$

It is actually two recurrence relations — one for even k, one for odd k. Thus, given the values of $a_0^{(n)}$ and $a_1^{(n)}$, we can work out all $a_k^{(n)}$, and so determine $\chi_n(q)$, and then have a solution of the differential equation. This solution depends on the value chosen for n, which is the looked-for eigenvalue that we should determine first.

To this end, let us assume we have found an eigenvalue n. Then

$$\begin{aligned} k = 2j \text{ even}: &\quad a_{2(j+1)}^{(n)} = 4\Big(j - \frac{n}{2}\Big)a_{2j}^{(n)} \,, \\ k = 2j+1 \text{ odd}: &\quad a_{2(j+1)+1}^{(n)} = 4\bigg(j - \frac{n-1}{2}\bigg)a_{2j+1}^{(n)} \,, \end{aligned} \tag{5.3.23}$$

for $j = 0, 1, 2, \ldots$. For large k, that is large j, these are approximately

$$\begin{aligned} a_{2(j+1)}^{(n)} &\cong 4j a_{2j}^{(n)} \cong 4(j+1)a_{2j}^{(n)} \,, \\ a_{2(j+1)+1}^{(n)} &\cong 4j a_{2j+1}^{(n)} \cong 4(j+1)a_{2j+1}^{(n)} \,, \end{aligned} \tag{5.3.24}$$

or

$$a_{2j}^{(n)} \propto 4^j j! \quad \text{and} \quad a_{2j+1}^{(n)} \propto 4^j j! \,. \tag{5.3.25}$$

So the even terms together would amount roughly to

$$\propto \sum_j \frac{1}{(2j)!} 4^j j!\, q^{2j} \tag{5.3.26}$$

and the odd terms to something similar. We focus on large j here, so that Stirling's* approximation

$$j! \cong \sqrt{2\pi j}\, j^j \, \mathrm{e}^{-j} \quad (\text{for } j \gg 1) \tag{5.3.27}$$

applies, giving us

$$\frac{4^j j!}{(2j)!} \cong \frac{4^j \sqrt{2\pi j}\, j^j \, \mathrm{e}^{-j}}{\sqrt{4\pi j}\, (2j)^{2j} \, \mathrm{e}^{-2j}} = \frac{1}{\sqrt{2}} j^{-j} \, \mathrm{e}^j \cong \frac{\sqrt{\pi j}}{j!} \,. \tag{5.3.28}$$

Accordingly,

$$\sum_j \frac{4^j j!}{(2j)!} q^{2j} \cong \sum_j \frac{\sqrt{\pi j}}{j!} q^{2j} \tag{5.3.29}$$

*James Stirling (1692–1770)

grows like e^{q^2} (in fact, even a bit faster because of the $\sqrt{j}$ in the numerator), so that the large-j behavior of $\chi_n(q)$ is dominated by e^{q^2} and that of $\psi_n(q)$,

$$\psi_n(q) = \chi_n(q)\,\mathrm{e}^{-\frac{1}{2}q^2} \sim \mathrm{e}^{\frac{1}{2}q^2}\,, \tag{5.3.30}$$

is then of the form that we had to discard earlier. The only way out of this dilemma is that the eigenvalue n must be an integer itself so that $k-n$ is zero for some k and then the recurrence relation stops. In other words, there are no $a_k^{(n)} \neq 0$ for arbitrary large k values, and therefore we do not have this problem.

In summary, then, we must have

$$\begin{aligned} n = \text{even integer}, \quad a_n^{(n)} \neq 0\,, \qquad & a_{2j}^{(n)} = 0 \text{ for } 2j > n\,, \\ & a_{2j+1}^{(n)} = 0 \text{ for all } j\,; \\ \text{or} \quad n = \text{odd integer}, \quad a_n^{(n)} \neq 0\,, \qquad & a_{2j+1}^{(n)} = 0 \text{ for } 2j+1 > n\,, \\ & a_{2j}^{(n)} = 0 \text{ for all } j\,. \end{aligned} \tag{5.3.31}$$

That is, $\chi_n(q)$ is a polynomial of degree n that is even in q if n is even and odd in q if n is odd. If we put, by matter of convention,

$$a_n^{(n)} = 2^n\,, \tag{5.3.32}$$

then $\chi_n(q)$ is the nth *Hermite polynomial* $\mathrm{H}_n(q)$. These polynomials can be defined, as we did, by their differential equation,

$$\left(\frac{\mathrm{d}^2}{\mathrm{d}q^2} - 2q\frac{\mathrm{d}}{\mathrm{d}q} + 2n\right)\mathrm{H}_n(q) = 0\,, \tag{5.3.33}$$

in conjunction with requiring that the highest power of q is

$$\mathrm{H}_n(q) = (2q)^n + \big(\text{powers } q^{n-2},\ q^{n-4},\ \ldots\big)\,. \tag{5.3.34}$$

The symmetry

$$\mathrm{H}_n(-q) = (-1)^n\mathrm{H}_n(q) \tag{5.3.35}$$

then follows.

The energy eigenvalues of the harmonic oscillator are thus

$$\begin{aligned} E_n &= \hbar\omega\left(n + \frac{1}{2}\right) \quad \text{with} \quad n = 0, 1, 2, \ldots \\ &= \frac{1}{2}\hbar\omega,\ \frac{3}{2}\hbar\omega,\ \frac{5}{2}\hbar\omega,\ \ldots \end{aligned} \tag{5.3.36}$$

and the corresponding wave functions are of the form

$$\psi_n(x) = a_n \mathrm{H}_n(q)\, \mathrm{e}^{-\frac{1}{2}q^2} \quad \text{with} \quad q = \sqrt{\frac{M\omega}{\hbar}}\, x \tag{5.3.37}$$

and a normalization factor a_n that is undetermined as yet. We shall get it in passing later; see (5.3.75) and (5.3.82). As a consequence of the symmetry property (5.3.35), the oscillator eigenstates are odd states when n is odd and even states when n is even.

Note that the ground-state energy of the Hamilton operator in (5.3.1),

$$H = \frac{1}{2M}P^2 + \frac{1}{2}M\omega^2 X^2\,, \tag{5.3.38}$$

is $\frac{1}{2}\hbar\omega$ (obtained for $n = 0$, of course). But we know that it is always possible to add an arbitrary constant to H without changing anything of relevance. Therefore,

$$H = \frac{1}{2M}P^2 + \frac{1}{2}M\omega^2 X^2 - \frac{1}{2}\hbar\omega \quad \text{with} \quad E_n = n\hbar\omega \tag{5.3.39}$$

would be just as good, and it is quite often more convenient to use this alternative Hamilton operator, for which the ground-state energy is $E_0 = 0$.

5.3.2 *Energy eigenstates: Ladder-operator approach*

Since we measure energies in units of $\hbar\omega$, we can exhibit this factor in H itself by writing

$$H = \hbar\omega\left(\frac{M\omega}{2\hbar}X^2 + \frac{1}{2\hbar M\omega}P^2\right). \tag{5.3.40}$$

This looks like the absolute square of a complex number with real part $\propto X$ and imaginary part $\propto P$, but we are dealing with noncommuting operators here and, therefore, the absolute square of

$$A = \sqrt{\frac{M\omega}{2\hbar}}X + \mathrm{i}\frac{1}{\sqrt{2\hbar M\omega}}P \tag{5.3.41}$$

requires multiplication with the adjoint operator,

$$A^\dagger = \sqrt{\frac{M\omega}{2\hbar}}X - \mathrm{i}\frac{1}{\sqrt{2\hbar M\omega}}P\,, \tag{5.3.42}$$

and the multiplication order matters. See

$$\left.\begin{matrix} AA^\dagger \\ A^\dagger A \end{matrix}\right\} = \frac{M\omega}{2\hbar}X^2 + \frac{1}{2\hbar M\omega}P^2 \mp \mathrm{i}\frac{1}{2\hbar}\underbrace{(XP-PX)}_{=\mathrm{i}\hbar}$$
$$= \frac{1}{\hbar\omega}H \pm \frac{1}{2}\,. \tag{5.3.43}$$

Accordingly, we have the commutation relation

$$[A, A^\dagger] = 1 \tag{5.3.44}$$

and can express the Hamilton operator in various ways, including these three versions:

$$H = \hbar\omega\left(A^\dagger A + \frac{1}{2}\right) = \hbar\omega\left(AA^\dagger - \frac{1}{2}\right)$$
$$= \frac{1}{2}\hbar\omega\big(A^\dagger A + AA^\dagger\big)\,, \tag{5.3.45}$$

the first of which is most commonly used.

The energy eigenstates $|n\rangle$,

$$H\big|n\big\rangle = \big|n\big\rangle\hbar\omega\left(n+\frac{1}{2}\right), \tag{5.3.46}$$

are therefore eigenstates of $A^\dagger A$,

$$A^\dagger A\big|n\big\rangle = \big|n\big\rangle n \quad \text{with} \quad n = 0, 1, 2, \dots\ . \tag{5.3.47}$$

In this context, it is common to refer to the states $|n\rangle$ as the *Fock* states*. In particular, we have

$$\big\langle n\big|A^\dagger A\big|n\big\rangle = n\,, \tag{5.3.48}$$

which tells us that the ket $A|n\rangle$ has length $\sqrt{n}$. For $n = 0$, it follows that

$$A\big|0\big\rangle = 0\,. \tag{5.3.49}$$

Therefore, the ground-state wave function $\psi_0(x) = \langle x|0\rangle$ obeys

$$0 = \big\langle x\big|A\big|0\big\rangle = \left(\sqrt{\frac{M\omega}{2\hbar}}\,x + \mathrm{i}\frac{1}{\sqrt{2\hbar M\omega}}\frac{\hbar}{\mathrm{i}}\frac{\partial}{\partial x}\right)\psi_0(x) \tag{5.3.50}$$

or

$$0 = \left(\frac{\partial}{\partial x} + \frac{M\omega}{\hbar}x\right)\psi_0(x)\,. \tag{5.3.51}$$

*Vladimir Alexandrovich FOCK (1898–1974)

This is (4.8.4) with

$$\frac{1}{2(\delta X)^2} = \frac{M\omega}{\hbar} \quad \text{or} \quad \delta X = \sqrt{\frac{\hbar}{2M\omega}} \tag{5.3.52}$$

and, therefore, $\psi_0(x)$ is the minimum-uncertainty wave function of (4.8.10),

$$\psi_0(x) = \left(\frac{M\omega}{\pi\hbar}\right)^{\frac{1}{4}} \mathrm{e}^{-\frac{1}{2}\frac{M\omega}{\hbar}x^2} . \tag{5.3.53}$$

The corresponding momentum wave function $\psi_0(p) = \langle p|0\rangle$ is available in (4.8.15) with

$$\delta P = \frac{\hbar/2}{\delta X} = \sqrt{\frac{1}{2}\hbar M\omega} \tag{5.3.54}$$

so that

$$\psi_0(p) = \big(\pi\hbar M\omega\big)^{-\frac{1}{4}} \mathrm{e}^{-\frac{1}{2}\frac{1}{\hbar M\omega}p^2} . \tag{5.3.55}$$

Now notice the following:

$$A^\dagger A A^\dagger = A^\dagger\big(\underbrace{AA^\dagger - A^\dagger A}_{=1} + A^\dagger A\big) = A^\dagger\big(A^\dagger A + 1\big) . \tag{5.3.56}$$

Apply it to $|n\rangle$,

$$A^\dagger A A^\dagger\big|n\big\rangle = A^\dagger\big(A^\dagger A + 1\big)\big|n\big\rangle = A^\dagger\big|n\big\rangle(n+1) , \tag{5.3.57}$$

and learn that $A^\dagger\big|n\big\rangle$ is the eigenket of $A^\dagger A$ with eigenvalue $n{+}1$. Therefore, it must be proportional to the eigenket $|n+1\rangle$,

$$A^\dagger\big|n\big\rangle \propto \big|n+1\big\rangle , \tag{5.3.58}$$

and

$$\begin{aligned}\Big(\big\langle n\big|A\Big)\Big(A^\dagger\big|n\big\rangle\Big) &= \big\langle n\big|AA^\dagger\big|n\big\rangle \\ &= \big\langle n\big|\big(A^\dagger A + 1\big)\big|n\big\rangle = n+1\end{aligned} \tag{5.3.59}$$

implies that

$$A^\dagger\big|n\big\rangle = \big|n+1\big\rangle\sqrt{n+1} \tag{5.3.60}$$

up to an arbitrary phase factor, which we put equal to 1, thereby adopting the standard conventions.

The adjoint statement is

$$\langle n|A = \sqrt{n+1}\,\langle n+1|\,. \tag{5.3.61}$$

We act with $A^\dagger$ on the right,

$$\begin{aligned}\langle n|AA^\dagger &= \langle n|\big(A^\dagger A + 1\big) = (n+1)\langle n| \\ &= \sqrt{n+1}\,\langle n+1|A^\dagger\,,\end{aligned} \tag{5.3.62}$$

and conclude that

$$\langle n+1|A^\dagger = \sqrt{n+1}\,\langle n|\,, \tag{5.3.63}$$

to which

$$A|n+1\rangle = |n\rangle\sqrt{n+1} \tag{5.3.64}$$

is the adjoint. To ease memorization, we collect all these statements,

$$\begin{aligned} A^\dagger|n\rangle &= |n+1\rangle\sqrt{n+1}\,, & \langle n|A &= \sqrt{n+1}\,\langle n+1|\,, \\ A|n\rangle &= |n-1\rangle\sqrt{n}\,, & \langle n|A^\dagger &= \sqrt{n}\,\langle n-1|\,. \end{aligned} \tag{5.3.65}$$

These relations show how A and $A^\dagger$ raise or lower the quantum number n by 1. In view of this property, they are called *ladder operators* — you go up or down the n ladder by one rung upon applying $A^\dagger$ or A to the bra or ket in question. It is customary to speak of the "raising operator $A^\dagger$" and the "lowering operator A," which is their action on kets $|n\rangle$, but keep in mind that the roles of raising and lowering operators are reversed for bras $\langle n|$.

This raising and lowering action by itself can be used to determine the eigenvalues of $A^\dagger A$, and thus of the Hamilton operator H, in a purely algebraic manner without resorting to a direct solution of the time-independent Schrödinger equation (5.3.4) or (5.3.11). So, let us pretend that we do not know as yet the eigenvalues of $A^\dagger A$ but agree to write

$$A^\dagger A|n\rangle = |n\rangle n \tag{5.3.66}$$

without any assumptions which real values n can take on. The ladder operator property is an application of the commutation relation $[A, A^\dagger] = 1$, so it is surely true that $A^\dagger|n\rangle = |n+1\rangle\sqrt{n+1}$ and so forth. This is to say that if n is an eigenvalue of $A^\dagger A$, then also $n+1$ is an eigenvalue of $A^\dagger A$, so we can climb up the ladder. Climbing down is done by applying A, $A|n\rangle = |n-1\rangle\sqrt{n}$, but that cannot go on forever because we know from

(5.3.48) that the squared length of $A|n\rangle$ equals n, which therefore cannot be negative.

So, when climbing down the ladder there must be a last rung, a bottom rung, for which the application of A does not give another eigenket of $A^\dagger A$. This is only possible if $n = 0$ marks the bottom rung because $A|n = 0\rangle = 0$. It follows that $n = 0$ is the smallest eigenvalue of $A^\dagger A$, and then $n = 1, 2, 3, \ldots$ are the other eigenvalues, obtained by moving up on the ladder. Together, we have

$$n = 0,\, 1,\, 2,\, 3,\, \ldots\,, \tag{5.3.67}$$

and we get this result without the harder work in Section 5.3.1, where we extracted this information out of the Schrödinger eigenvalue equation for $A^\dagger A = H/\hbar\omega - \frac{1}{2}$.

In fact, we are further rewarded by the algebraic method because we are told that

$$\begin{aligned} \big|n\big\rangle &= \frac{A^\dagger}{\sqrt{n}}\big|n-1\big\rangle = \frac{A^\dagger}{\sqrt{n}}\frac{A^\dagger}{\sqrt{n-1}}\big|n-2\big\rangle \\ &= \cdots = \frac{(A^\dagger)^n}{\sqrt{n!}}\big|0\big\rangle\,, \end{aligned} \tag{5.3.68}$$

that is, we obtain the eigenkets of $A^\dagger A$ together with their normalization. Beginning with

$$\psi_0(x) = \big\langle x\big|0\big\rangle = \left(\frac{M\omega}{\pi\hbar}\right)^{\frac{1}{4}} \mathrm{e}^{-\frac{1}{2}\frac{M\omega}{\hbar}x^2} = \left(\frac{M\omega}{\pi\hbar}\right)^{\frac{1}{4}} \mathrm{e}^{-\frac{1}{2}q^2}\,, \tag{5.3.69}$$

we can thus get the wave functions for all eigenstates of the harmonic oscillator by repeated application of $A^\dagger$,

$$\psi_n(x) = \big\langle x\big|n\big\rangle = \frac{1}{\sqrt{n!}}\big\langle x\big|\big(A^\dagger\big)^n\big|0\big\rangle \tag{5.3.70}$$

with

$$\begin{aligned} \big\langle x\big|A^\dagger &= \big\langle x\big|\left(\sqrt{\frac{M\omega}{2\hbar}}X - \mathrm{i}\frac{1}{\sqrt{2M\hbar\omega}}P\right) \\ &= \left(\sqrt{\frac{M\omega}{2\hbar}}x - \mathrm{i}\frac{1}{\sqrt{2M\hbar\omega}}\frac{\hbar}{\mathrm{i}}\frac{\partial}{\partial x}\right)\big\langle x\big| \\ &= \frac{1}{\sqrt{2}}\left(q - \frac{\partial}{\partial q}\right)\big\langle x\big|\,. \end{aligned} \tag{5.3.71}$$

Accordingly,

$$\psi_n(x) = \left(\frac{M\omega}{\pi\hbar}\right)^{\frac{1}{4}} \frac{1}{\sqrt{n!}} \left(\frac{1}{\sqrt{2}}\left(q - \frac{\mathrm{d}}{\mathrm{d}q}\right)\right)^n \mathrm{e}^{-\frac{1}{2}q^2}$$
$$= \left(\frac{M\omega}{\pi\hbar}\right)^{\frac{1}{4}} \frac{1}{\sqrt{2^n n!}} \left(q - \frac{\mathrm{d}}{\mathrm{d}q}\right)^n \mathrm{e}^{-\frac{1}{2}q^2} . \tag{5.3.72}$$

This becomes simpler once we note that

$$q - \frac{\mathrm{d}}{\mathrm{d}q} = \mathrm{e}^{\frac{1}{2}q^2} \left(-\frac{\mathrm{d}}{\mathrm{d}q}\right) \mathrm{e}^{-\frac{1}{2}q^2} \tag{5.3.73}$$

and therefore

$$\left(q - \frac{\mathrm{d}}{\mathrm{d}q}\right)^n = \mathrm{e}^{\frac{1}{2}q^2} \left(-\frac{\mathrm{d}}{\mathrm{d}q}\right)^n \mathrm{e}^{-\frac{1}{2}q^2} . \tag{5.3.74}$$

We have thus arrived at

$$\psi_n(x) = \left(\frac{M\omega}{\pi\hbar}\right)^{\frac{1}{4}} \frac{1}{\sqrt{2^n n!}} \mathrm{e}^{\frac{1}{2}q^2} \left(-\frac{\mathrm{d}}{\mathrm{d}q}\right)^n \mathrm{e}^{-q^2} , \tag{5.3.75}$$

the position wave function of the nth Fock state, which we compare with the earlier result (5.3.37) and infer that

$$\mathrm{H}_n(q) = \mathrm{e}^{q^2} \left(-\frac{\mathrm{d}}{\mathrm{d}q}\right)^n \mathrm{e}^{-q^2} . \tag{5.3.76}$$

There is no room here for a multiplicative constant; there is no extra factor. The normalization of $\mathrm{H}_n(q)$ in accordance with (5.3.34) is exactly what (5.3.76) gives, because the leading power is obtained by applying each $-\frac{\mathrm{d}}{\mathrm{d}q}$ to e^{-q^2}, each application supplying one factor of $2q$.

5.3.3 *Hermite polynomials*

There are very many mathematical identities, recurrence relations and others, for the Hermite polynomials, and quite a number of equivalent definitions for them. We have met two: the differential equation (5.3.11) and the *Rodrigues* formula* (5.3.76). Arguably, the simplest and also the most important formula is the *generating function*

$$\mathrm{e}^{2tq - t^2} = \sum_{n=0}^{\infty} \frac{t^n}{n!} \mathrm{H}_n(q) , \tag{5.3.77}$$

*Benjamin Olinde RODRIGUES (1795–1851)

from which all statements about $\mathrm{H}_n(q)$ can be derived rather easily. This generating function can be found by going backward,

$$\begin{aligned}\sum_{n=0}^{\infty}\frac{t^n}{n!}\mathrm{H}_n(q) &= \mathrm{e}^{q^2}\sum_{n=0}^{\infty}\frac{t^n}{n!}\left(-\frac{\mathrm{d}}{\mathrm{d}q}\right)^n\mathrm{e}^{-q^2}\\ &= \mathrm{e}^{q^2}\,\mathrm{e}^{-t\frac{\mathrm{d}}{\mathrm{d}q}}\,\mathrm{e}^{-q^2} = \mathrm{e}^{q^2}\,\mathrm{e}^{-(q-t)^2} = \mathrm{e}^{2qt-t^2}\,,\end{aligned}\tag{5.3.78}$$

where a variant of (4.3.3) is used.

As a simple application, we consider the orthogonality relation of the Hermite polynomials,

$$\int_{-\infty}^{\infty}\mathrm{d}q\;\mathrm{e}^{-q^2}\,\mathrm{H}_n(q)\,\mathrm{H}_m(q) = \boxed{?}_{mn}\,.\tag{5.3.79}$$

We deal with these numbers as a set: multiply by $\dfrac{t^n}{n!}\dfrac{s^m}{m!}$ and sum over n and m,

$$\begin{aligned}\sum_{n,m=0}^{\infty}\frac{t^n}{n!}\frac{s^m}{m!}\boxed{?}_{nm} &= \int_{-\infty}^{\infty}\mathrm{d}q\;\mathrm{e}^{-q^2}\,\mathrm{e}^{2qt-t^2}\,\mathrm{e}^{2sq-s^2}\\ &= \int\mathrm{d}q\;\mathrm{e}^{-\left(q-(t+s)\right)^2}\,\mathrm{e}^{2ts} = \sqrt{\pi}\,\mathrm{e}^{2ts}\\ &= \sqrt{\pi}\sum_{n=0}^{\infty}\frac{(2ts)^n}{n!} = \sqrt{\pi}\sum_{n,m=0}^{\infty}\frac{2^n t^n s^m}{n!}\delta_{nm}\end{aligned}\tag{5.3.80}$$

so that $\boxed{?}_{nm} = 2^n n!\sqrt{\pi}\,\delta_{nm}$ and

$$\int_{-\infty}^{\infty}\mathrm{d}q\;\mathrm{e}^{-q^2}\,\mathrm{H}_n(q)\,\mathrm{H}_m(q) = \delta_{nm}\,2^n n!\sqrt{\pi}\,.\tag{5.3.81}$$

In the physical context of eigenstates of the harmonic oscillator, this is essentially the orthonormality of the Fock state kets $|n\rangle$, explicitly specified by their position wave functions

$$\psi_n(x) = \left(\frac{M\omega}{\pi\hbar}\right)^{\frac{1}{4}}\left(2^n n!\right)^{-\frac{1}{2}}\mathrm{H}_n(q)\,\mathrm{e}^{-\frac{1}{2}q^2}\quad\text{with}\quad q=\sqrt{\frac{M\omega}{\hbar}}x\,.\tag{5.3.82}$$

We check this by inserting $\psi_n(x)$ and $\psi_m(x)$ into

$$\langle n|m\rangle = \int\mathrm{d}x\,\langle n|x\rangle\langle x|m\rangle = \int\mathrm{d}x\,\psi_n(x)^*\psi_m(x)\tag{5.3.83}$$

and find

$$\langle n|m\rangle = \left(\frac{M\omega}{\pi\hbar}\right)^{\frac{1}{2}} \left(2^{n+m}\, n!\, m!\right)^{-\frac{1}{2}} \int \underbrace{\mathrm{d}x}_{=\sqrt{\hbar/(M\omega)}\,\mathrm{d}q} \mathrm{e}^{-q^2} \mathrm{H}_n(q) \mathrm{H}_m(q)$$
$$= \left(2^{n+m}\, n!\, m!\, \pi\right)^{-\frac{1}{2}} \delta_{nm} 2^n n! \sqrt{\pi} = \delta_{nm}\,, \tag{5.3.84}$$

indeed.

5.3.4 *Infinite matrices*

When using the oscillator eigenkets $|0\rangle, |1\rangle, \ldots$ as the basis kets and their bras as the basis bras, an arbitrary operator Z is represented by an *infinite* matrix,

$$Z = \sum_{n=0}^{\infty} |n\rangle\langle n|Z \sum_{m=0}^{\infty} |m\rangle\langle m| = \sum_{n,m=0}^{\infty} |n\rangle\langle n|Z|m\rangle\langle m|$$
$$= \Big(|0\rangle\ |1\rangle \cdots \Big) \begin{pmatrix} \langle 0|Z|0\rangle & \langle 0|Z|1\rangle & \cdots \\ \langle 1|Z|0\rangle & \langle 1|Z|1\rangle & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix} \begin{pmatrix} \langle 0| \\ \langle 1| \\ \vdots \end{pmatrix}$$
$$\widehat{=} \begin{pmatrix} \langle 0|Z|0\rangle & \langle 0|Z|1\rangle & \cdots \\ \langle 1|Z|0\rangle & \langle 1|Z|1\rangle & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix}, \tag{5.3.85}$$

and kets are represented by infinitely long columns,

$$|\ \rangle = \Big(|0\rangle\ |1\rangle \cdots \Big) \begin{pmatrix} \langle 0|\ \rangle \\ \langle 1|\ \rangle \\ \vdots \end{pmatrix} \widehat{=} \begin{pmatrix} \langle 0|\ \rangle \\ \langle 1|\ \rangle \\ \vdots \end{pmatrix} = \begin{pmatrix} \psi_0 \\ \psi_1 \\ \vdots \end{pmatrix}, \tag{5.3.86}$$

and bras by infinitely long rows. Matrix multiplication, or the multiplication of a matrix with a column, then involves an infinite summation, a series, and the convergence properties offer the usual treat of possible pitfalls, easily avoided as a rule by a bit of careful attention to such details.

As illustrating examples, let us find the infinite matrices for the position operator X and the momentum operator P. We recall (5.3.41) and (5.3.42) and note that

$$X = \sqrt{\frac{\hbar}{2M\omega}} \left(A^\dagger + A\right), \quad P = \sqrt{\frac{\hbar M\omega}{2}} \left(\mathrm{i}A^\dagger - \mathrm{i}A\right). \tag{5.3.87}$$

Therefore, we have

$$\begin{aligned}\langle n|X|m\rangle &= \sqrt{\frac{\hbar}{2M\omega}}\,\langle n|(A^\dagger + A)|m\rangle \\ &= \sqrt{\frac{\hbar}{2M\omega}}\Big(\sqrt{n}\,\langle n-1|m\rangle + \langle n|m-1\rangle\sqrt{m}\Big) \\ &= \sqrt{\frac{\hbar}{2M\omega}}\,\Big(\sqrt{n}\,\delta_{n,m+1} + \sqrt{m}\,\delta_{n+1,m}\Big) \end{aligned} \tag{5.3.88}$$

and likewise,

$$\langle n|P|m\rangle = \sqrt{\frac{\hbar M\omega}{2}}\,\Big(\mathrm{i}\sqrt{n}\,\delta_{n,m+1} - \mathrm{i}\sqrt{m}\,\delta_{n+1,m}\Big)\,. \tag{5.3.89}$$

Quite explicitly, the matrices are

$$X \mathrel{\widehat{=}} \sqrt{\frac{\hbar}{2M\omega}}\begin{pmatrix} 0 & \sqrt{1} & 0 & \cdots & \cdots & \cdots \\ \sqrt{1} & 0 & \sqrt{2} & 0 & \cdots & \cdots \\ 0 & \sqrt{2} & 0 & \sqrt{3} & 0 & \cdots \\ 0 & 0 & \sqrt{3} & 0 & \sqrt{4} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} \tag{5.3.90}$$

and

$$P \mathrel{\widehat{=}} \sqrt{\frac{\hbar M\omega}{2}}\begin{pmatrix} 0 & -\mathrm{i}\sqrt{1} & 0 & \cdots & \cdots & \cdots \\ \mathrm{i}\sqrt{1} & 0 & -\mathrm{i}\sqrt{2} & 0 & \cdots & \cdots \\ 0 & \mathrm{i}\sqrt{2} & 0 & -\mathrm{i}\sqrt{3} & 0 & \cdots \\ 0 & 0 & \mathrm{i}\sqrt{3} & 0 & -\mathrm{i}\sqrt{4} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}, \tag{5.3.91}$$

and it is an easy task to verify that

$$XP - PX = [X,P] \mathrel{\widehat{=}} \mathrm{i}\hbar\begin{pmatrix} 1 & 0 & \cdots & \cdots & \cdots \\ 0 & 1 & 0 & \cdots & \cdots \\ 0 & 0 & 1 & 0 & \cdots \\ \vdots & \vdots & \vdots & \ddots & \ddots \end{pmatrix}, \tag{5.3.92}$$

where we have an infinite identity matrix.

5.3.5 *Position and momentum spreads*

We know from the discussion in Section 5.3.2 that the ground state ($n = 0$) of the harmonic oscillator is a minimum-uncertainty state. For arbitrary n, we first note that

$$\begin{aligned}\langle X\rangle_n &= \sqrt{\frac{\hbar}{2M\omega}}\,\langle n|\big(A^\dagger + A\big)|n\rangle = 0\,,\\ \langle P\rangle_n &= \sqrt{\frac{\hbar M\omega}{2}}\,\langle n|\big(\mathrm{i}A^\dagger - \mathrm{i}A\big)|n\rangle = 0\,,\end{aligned} \tag{5.3.93}$$

and further we have

$$\begin{aligned}\Big\langle X^2\Big\rangle_n &= \frac{\hbar}{2M\omega}\,\langle n|\big(A^\dagger + A\big)^2|n\rangle\\ &= \frac{\hbar}{2M\omega}\,\langle n|\Big(\underset{\downarrow}{A^{\dagger 2}} + \underset{\downarrow}{A^\dagger A} + \underset{\downarrow}{AA^\dagger} + \underset{\downarrow}{A^2}\Big)|n\rangle\\ &\qquad\qquad\quad 0 \qquad\quad n \qquad\quad n+1 \qquad\quad 0\\ &= \frac{\hbar}{M\omega}\left(n + \frac{1}{2}\right)\end{aligned} \tag{5.3.94}$$

and

$$\begin{aligned}\Big\langle P^2\Big\rangle_n &= \frac{\hbar M\omega}{2}\langle n|\big(\mathrm{i}A^\dagger - \mathrm{i}A\big)^2|n\rangle\\ &= \frac{\hbar M\omega}{2}\langle n|\Big(-\underset{\downarrow}{A^{\dagger 2}} + \underset{\downarrow}{A^\dagger A} + \underset{\downarrow}{AA^\dagger} - \underset{\downarrow}{A^2}\Big)|n\rangle\\ &\qquad\qquad\quad 0 \qquad\quad n \qquad\quad n+1 \qquad\quad 0\\ &= \hbar M\omega\left(n + \frac{1}{2}\right).\end{aligned} \tag{5.3.95}$$

Accordingly, we find

$$\begin{aligned}\delta X_n &= \sqrt{\frac{\hbar}{M\omega}}\,\sqrt{n + \frac{1}{2}}\,,\\ \delta P_n &= \sqrt{\hbar M\omega}\,\sqrt{n + \frac{1}{2}}\,,\end{aligned} \tag{5.3.96}$$

for the position spread and momentum spread in the nth oscillator state. Their products

$$\delta X_n\,\delta P_n = \hbar\left(n + \frac{1}{2}\right) = \frac{1}{2}\hbar,\ \frac{3}{2}\hbar,\ \frac{5}{2}\hbar,\ \ldots \tag{5.3.97}$$

always exceed Heisenberg's lower bound of $\frac{1}{2}\hbar$ when $n > 0$.

5.4 Delta potential

5.4.1 *Bound state*

On a couple of occasions — see, for instance, Exercises 76 and 81 — we used the wave function

$$\psi(x) = \sqrt{\kappa}\, \mathrm{e}^{-\kappa|x|}, \quad \kappa > 0 \tag{5.4.1}$$

as an example. Its derivative

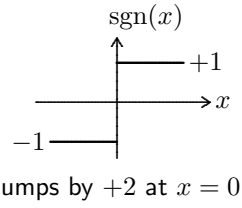
$$\frac{\partial}{\partial x}\psi(x) = -\sqrt{\kappa^3}\, \mathrm{sgn}(x)\, \mathrm{e}^{-\kappa|x|} \tag{5.4.2}$$

involves the derivative of $|x|$,

$$\frac{\mathrm{d}}{\mathrm{d}x}|x| = \mathrm{sgn}(x) = \begin{cases} +1 & \text{for} \quad x > 0\,, \\ -1 & \text{for} \quad x < 0\,, \end{cases} \tag{5.4.3}$$

the sign function that we first met in (5.2.24). The second derivative of $\psi(x)$ needs the derivative thereof, see Exercise 108,

$$\frac{\mathrm{d}}{\mathrm{d}x}\,\mathrm{sgn}(x) = 2\,\delta(x) \tag{5.4.4}$$



so that

$$\begin{aligned} \left(\frac{\partial}{\partial x}\right)^2 \psi(x) &= \sqrt{\kappa^5}\, \mathrm{e}^{-\kappa|x|} - 2\sqrt{\kappa^3}\, \delta(x)\, \mathrm{e}^{-\kappa|x|} \\ &= \kappa^2 \psi(x) - 2\kappa\, \delta(x)\psi(x)\,. \end{aligned} \tag{5.4.5}$$

We compare (5.4.5) with the Schrödinger eigenvalue equation for a Hamilton operator with potential energy $V(X)$,

$$H = \frac{1}{2M}P^2 + V(X)\,, \tag{5.4.6}$$

that is,

$$-\frac{\hbar^2}{2M}\left(\frac{\partial}{\partial x}\right)^2 \psi_E(x) + V(x)\psi_E(x) = E\psi_E(x) \tag{5.4.7}$$

or

$$\left(\frac{\partial}{\partial x}\right)^2 \psi_E(x) = -\frac{2ME}{\hbar^2}\psi_E(x) + \frac{2M}{\hbar^2}V(x)\psi_E(x) \tag{5.4.8}$$

and observe that $\psi(x) = \sqrt{\kappa}\,\mathrm{e}^{-\kappa|x|}$ solves this equation for

$$\kappa^2 = -\frac{2ME}{\hbar^2} \quad \text{and} \quad -2\kappa\,\delta(x) = \frac{2M}{\hbar^2}V(x) \tag{5.4.9}$$

or

$$V(x) = -\frac{\hbar^2\kappa}{M}\,\delta(x) \quad \text{and} \quad E = -\frac{(\hbar\kappa)^2}{2M}\,. \tag{5.4.10}$$

The Hamilton operator in question is, therefore,

$$H = \frac{1}{2M}P^2 - \frac{\hbar^2\kappa}{M}\delta(X)\,. \tag{5.4.11}$$

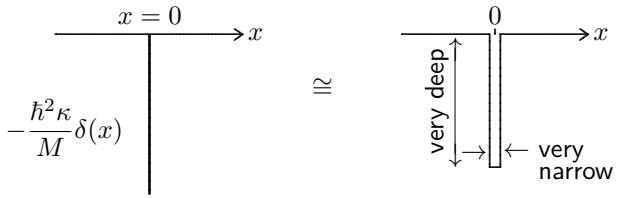
The Dirac delta function of position operator X that appears here has a rather simple meaning, inasmuch as

$$\begin{aligned}\delta(X) &= \int \mathrm{d}x\,|x\rangle\,\delta(x)\,\langle x| \\ &= \Big(|x\rangle\langle x|\Big)_{x\,=\,0} = |x=0\rangle\langle x=0|\,;\end{aligned} \tag{5.4.12}$$

it is the projector to the state for the eigenvalue $x = 0$ of the operator X. Geometrically speaking, the potential energy

$$V(x) = -\frac{\hbar^2\kappa}{M}\,\delta(x) \tag{5.4.13}$$

approximates the more physical situation of a very narrow and very deep potential well:



(5.4.14)

or, put differently, a very strong force of very short range.

The modeling of the more realistic physical situation by the mathematical idealization of a delta potential is justifiable if all relevant distances are large compared with the range of the force. In particular, the distance

over which the wave function of interest changes significantly must be large in this sense. If this is the case, the simple $\delta(x)$ potential catches the essence of the strong short-range force well and we use it as a good physical approximation to the more complicated real thing.

Suppose we had not known the solution to the eigenvalue equation to begin with but have to find $|E\rangle$ such that

$$\left(\frac{1}{2M}P^2 - \frac{\hbar^2\kappa}{M}\delta(X)\right)\big|E\big\rangle = \big|E\big\rangle E\,. \tag{5.4.15}$$

How could we go about it? We could, of course, turn this into the differential equation for $\langle x|E\rangle$ and then solve it as it comes. It is, however, simpler and also instructive to look for the momentum wave function $\psi_E(p) = \langle p|E\rangle$ first.

In view of (5.4.12), we have

$$\begin{aligned}\big\langle p\big|\delta(X)\big|E\big\rangle &= \Big(\big\langle p\big|x\big\rangle\big\langle x\big|E\big\rangle\Big)_{x\,=\,0} \\ &= \left(\frac{\mathrm{e}^{-\mathrm{i}px/\hbar}}{\sqrt{2\pi\hbar}}\psi_E(x)\right)_{x\,=\,0} = \frac{1}{\sqrt{2\pi\hbar}}\psi_E(x=0)\,,\end{aligned} \tag{5.4.16}$$

where $\psi_E(x=0)$ is a number to be determined. Then, (5.4.15) turns into

$$\frac{1}{2M}p^2\psi_E(p) - \frac{\hbar^2\kappa}{M}\frac{\psi_E(x=0)}{\sqrt{2\pi\hbar}} = E\psi_E(p)\,. \tag{5.4.17}$$

Upon writing

$$\frac{\hbar^2\kappa}{M}\frac{\psi_E(x=0)}{\sqrt{2\pi\hbar}} = \frac{C}{2M} \quad \text{and} \quad E = -\frac{q^2}{2M} \quad \text{with} \quad q > 0\,, \tag{5.4.18}$$

this appears as

$$\big(p^2 + q^2\big)\psi_E(p) = C\,. \tag{5.4.19}$$

The solution

$$\psi_E(p) = \frac{C}{p^2 + q^2} \tag{5.4.20}$$

will give us also the position wave function by Fourier transformation,

$$\psi_E(x) = \int \mathrm{d}p\,\big\langle x\big|p\big\rangle\,\underbrace{\big\langle p\big|E\big\rangle}_{=\,\psi_E(p)} = \int \mathrm{d}p\,\frac{\mathrm{e}^{\mathrm{i}xp/\hbar}}{\sqrt{2\pi\hbar}}\frac{C}{p^2+q^2}\,. \tag{5.4.21}$$

But before we evaluate this integral for arbitrary x, let us look at the consistency condition that obtains for $x = 0$,

$$\psi_E(x=0) = \int \mathrm{d}p \, \frac{C}{\sqrt{2\pi\hbar}} \frac{1}{p^2+q^2} = \frac{C}{\sqrt{2\pi\hbar}} \underbrace{\int \frac{\mathrm{d}p}{p^2+q^2}}_{=\pi/q} = \frac{\hbar\kappa}{q} \psi_E(x=0)\,. \tag{5.4.22}$$

This tells us that $q = \hbar\kappa$ is the only possible value for q and that, therefore, the only negative eigenvalue of H is that of (5.4.10), $E = -(\hbar\kappa)^2/(2M)$.

We thus have

$$\psi_E(p) = \frac{C}{p^2 + (\hbar\kappa)^2} \tag{5.4.23}$$

and can determine the modulus of C, and thus of $\psi_E(x=0)$, from the normalization $\langle E|E\rangle = 1$,

$$1 = \int \mathrm{d}p \, \big|\psi_E(p)\big|^2 = \big|C\big|^2 \underbrace{\int \frac{\mathrm{d}p}{(p^2+q^2)^2}}_{=\pi/(2q^3)} \quad \text{with} \quad q = \hbar\kappa \tag{5.4.24}$$

so that

$$\big|C\big| = \sqrt{\frac{2}{\pi}}\,(\hbar\kappa)^{\frac{3}{2}} \quad \text{and} \quad \big|\psi_E(x=0)\big| = \sqrt{\kappa}\,, \tag{5.4.25}$$

consistent, of course, with (5.4.1). When opting for the phase choice specified by $\psi_E(x=0) > 0$, we get

$$\psi_E(x=0) = \sqrt{\kappa} \tag{5.4.26}$$

and

$$\psi_E(p) = \frac{\sqrt{2/\pi}(\hbar\kappa)^{\frac{3}{2}}}{p^2 + (\hbar\kappa)^2}\,. \tag{5.4.27}$$

We know from Exercise 81 that this is the momentum wave function for the position wave function in (5.4.1).

As a check, we return to the Fourier integral in (5.4.21) and evaluate it by a contour integration and an application of the residue method. You may skip ahead to Section 5.4.2 if you do not know enough complex analysis.

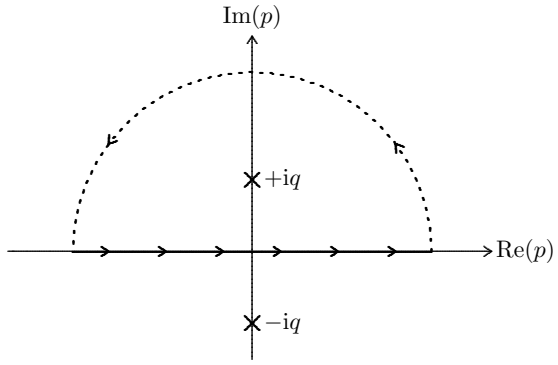
The integral in (5.4.21) defines a function that is even in x because the odd part of

$$\mathrm{e}^{\mathrm{i}xp/\hbar} = \underbrace{\cos(xp/\hbar)}_{\text{even}} + \underbrace{\mathrm{i}\sin(xp/\hbar)}_{\text{odd}} \tag{5.4.28}$$

does not contribute to the p integral. We can therefore write

$$\begin{aligned}\psi_E(x) &= \frac{C}{\sqrt{2\pi\hbar}} \int \mathrm{d}p\, \frac{\mathrm{e}^{\mathrm{i}|x|p/\hbar}}{p^2+q^2} \\ &= \frac{C}{\sqrt{2\pi\hbar}} \frac{1}{2\mathrm{i}q} \int \mathrm{d}p \left(\frac{\mathrm{e}^{\mathrm{i}|x|p/\hbar}}{p-\mathrm{i}q} - \frac{\mathrm{e}^{\mathrm{i}|x|p/\hbar}}{p+\mathrm{i}q} \right).\end{aligned} \tag{5.4.29}$$

The latter form exhibits the poles at $p = \mathrm{i}q$ and $p = -\mathrm{i}q$,



(5.4.30)

and we close the integration path by a very large half circle in the upper plane where $\mathrm{Im}(p) > 0$ and $\mathrm{e}^{\mathrm{i}|x|p/\hbar}$ is exponentially small. The p integral is therefore given by the residue at $p = +\mathrm{i}q$,

$$\begin{aligned}\psi_E(x) &= \frac{C}{\sqrt{2\pi\hbar}} \frac{1}{2\mathrm{i}q} 2\pi\mathrm{i} \underbrace{\mathrm{e}^{\mathrm{i}|x|\mathrm{i}q/\hbar}}_{\text{residue at } p=+\mathrm{i}q} \\ &= \sqrt{\frac{\pi}{2\hbar}} \frac{C}{q}\, \mathrm{e}^{-|x|q/\hbar} = \sqrt{\kappa}\, \mathrm{e}^{-\kappa|x|}\,,\end{aligned} \tag{5.4.31}$$

where $q = \hbar\kappa$ and $C = \sqrt{2\hbar\kappa/\pi}\,\hbar\kappa$ enter in the last step, as required by (5.4.22) and (5.4.27). Indeed, we are back at our starting point (5.4.1).

5.4.2 *Scattering states*

Having thus established that there is one and only one negative-energy eigenvalue of

$$H = \frac{1}{2M}P^2 - \frac{\hbar^2\kappa}{M}\delta(X) \tag{5.4.32}$$

with energy $E = -(\hbar\kappa)^2/(2M)$, all other eigenstates must have positive energy, $E = (\hbar k)^2/(2M)$ with $k > 0$. For the force-free motion, we had even states, $\psi(x) \propto \cos(kx)$, and odd states, $\psi(x) \propto \sin(kx)$. The latter wave functions vanish at $x = 0$ and, therefore, these remain solutions of

$$\left(-\frac{\hbar^2}{2M}\frac{\partial^2}{\partial x^2} - \frac{\hbar^2\kappa}{M}\delta(x)\right)\psi(x) = \frac{(\hbar k)^2}{2M}\psi(x)\,. \tag{5.4.33}$$

So, odd solutions are

$$\psi_{\text{odd}}(x) = \sin(kx)\,, \tag{5.4.34}$$

where we do not care about the proper normalization (for the sake of simplicity only — we could care if we wanted to).

The even solutions will be of the form

$$\psi_{\text{even}}(x) = \cos\bigl(k|x| - \alpha\bigr)\,, \tag{5.4.35}$$

with a phase shift α with $-\frac{1}{2}\pi < \alpha < \frac{1}{2}\pi$ that we need to determine. It is clear that this is the generic form of the even wave functions because for $x \neq 0$, we must have some linear combinations of $\cos(kx)$ and $\sin(kx)$ or, equivalently, one of them with a shifted argument. Upon enforcing $\psi(-x) = \psi(x)$, the stated form is implied.

We need the second derivative of $\psi_{\text{even}}(x)$,

$$\begin{aligned}\left(\frac{\partial}{\partial x}\right)^2\psi_{\text{even}}(x) &= \frac{\mathrm{d}}{\mathrm{d}x}\Bigl(-k\,\mathrm{sgn}(x)\sin\bigl(k|x| - \alpha\bigr)\Bigr)\\ &= -k^2\cos\bigl(k|x| - \alpha\bigr) - 2k\,\delta(x)\sin\bigl(k|x| - \alpha\bigr)\\ &= -k^2\psi_{\text{even}}(x) + 2k\,\delta(x)\sin(\alpha)\,,\end{aligned} \tag{5.4.36}$$

where the observation of Exercise 102 is used. We insert it into the eigenvalue differential equation (5.4.33),

$$\frac{\partial^2}{\partial x^2}\psi_{\text{even}}(x) + 2\kappa\,\delta(x)\underbrace{\psi_{\text{even}}(x)}_{\to\,\psi_{\text{even}}(0) = \cos(\alpha)} = -k^2\psi_{\text{even}}(x)\,, \tag{5.4.37}$$

and get

$$k\sin(\alpha)+\kappa\cos(\alpha)=0 \tag{5.4.38}$$

or

$$\tan(\alpha)=-\frac{\kappa}{k}\,, \tag{5.4.39}$$

so that

$$\begin{aligned}\kappa=0:&\quad \alpha=0\,,\\ k\gg\kappa:&\quad \alpha\cong-\frac{\kappa}{k}\,,\\ k\ll\kappa:&\quad \alpha\cong-\frac{\pi}{2}\,.\end{aligned} \tag{5.4.40}$$

In particular, the $\alpha=0$ value for $\kappa=0$ is expected because we must recover the force-free solution $\psi_{\text{even}}(x)=\cos(kx)=\cos\bigl(k|x|\bigr)$ for $\kappa=0$.

For $E>0$, we thus have

$$\text{odd solutions}\quad \psi_{\text{odd}}(x)=\sin(kx) \tag{5.4.41}$$

and

$$\text{even solutions}\quad \psi_{\text{even}}(x)=\cos\bigl(k|x|-\alpha\bigr) \tag{5.4.42}$$

with $k=\sqrt{2mE/\hbar^2}>0$ and $\alpha=-\tan^{-1}(\kappa/k)$. Now, we ask the following question: Which linear combination of them,

$$\psi(x)=A\cos\bigl(k|x|-\alpha\bigr)+B\sin(kx)\,, \tag{5.4.43}$$

is such that

$$\psi(x)=\mathrm{e}^{\mathrm{i}kx}\quad\text{for}\quad x>0\,, \tag{5.4.44}$$

and what is then $\psi(x)$ for $x<0$? We have, for $x>0$,

$$\begin{aligned}\mathrm{e}^{\mathrm{i}kx}&=A\bigl(\cos(kx)\cos(\alpha)+\sin(kx)\sin(\alpha)\bigr)+B\sin(kx)\\ &=A\cos(\alpha)\cos(kx)+\bigl(A\sin(\alpha)+B\bigr)\sin(kx)\end{aligned} \tag{5.4.45}$$

and thus need

$$A\cos(\alpha)=1\,,\quad A\sin(\alpha)+B=\mathrm{i} \tag{5.4.46}$$

or

$$\begin{aligned}A&=\cos(\alpha)^{-1}=\sqrt{1+\tan(\alpha)^2}=\sqrt{1+(\kappa/k)^2}\,,\\ B&=\mathrm{i}-A\sin(\alpha)=\mathrm{i}-\tan(\alpha)=\mathrm{i}+\kappa/k\,.\end{aligned} \tag{5.4.47}$$

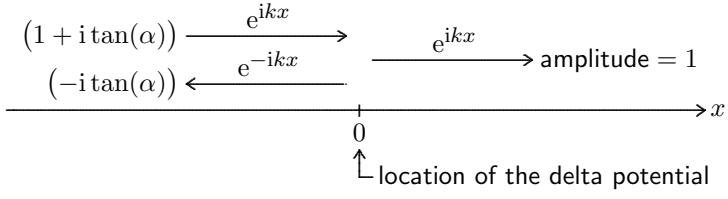
For $x < 0$, we obtain

$$\begin{aligned}\psi(x) &= A\cos(kx+\alpha) + B\sin(kx)\\ &= \frac{1}{\cos(\alpha)}\big(\cos(kx)\cos(\alpha) - \sin(kx)\sin(\alpha)\big) + \big(\mathrm{i} - \tan(\alpha)\big)\sin(kx)\\ &= \cos(kx) + \big(\mathrm{i} - 2\tan(\alpha)\big)\sin(kx)\\ &= \big(1 + \mathrm{i}\tan(\alpha)\big)\,\mathrm{e}^{\mathrm{i}kx} - \mathrm{i}\tan(\alpha)\,\mathrm{e}^{-\mathrm{i}kx}\,. \end{aligned} \tag{5.4.48}$$

In summary, this position wave function of an eigenstate to the energy $E = (\hbar k)^2/(2M)$ is of the form

$$\psi(x) = \begin{cases} \big(1 + \mathrm{i}\tan(\alpha)\big)\,\mathrm{e}^{\mathrm{i}kx} - \mathrm{i}\tan(\alpha)\,\mathrm{e}^{-\mathrm{i}kx} & \text{for} \quad x < 0\,, \\ \mathrm{e}^{\mathrm{i}kx} & \text{for} \quad x > 0\,. \end{cases} \tag{5.4.49}$$

Recalling that $P \to \dfrac{\hbar}{\mathrm{i}}\dfrac{\partial}{\partial x}$ for position wave functions, we have momentum $+\hbar k$ for $\mathrm{e}^{\mathrm{i}kx}$ and momentum $-\hbar k$ for $\mathrm{e}^{-\mathrm{i}kx}$, which tells us that $\mathrm{e}^{\mathrm{i}kx}$ is the amplitude for motion to the right (from negative to positive x values) and $\mathrm{e}^{-\mathrm{i}kx}$ for motion to the left:



(5.4.50)

We have here a particularly simple case of the more general situation discussed in Section 3.2 of *Perturbed Evolution*.

The wave function (5.4.49) is not normalized so that the amplitude factors that multiply the plane wave exponential factors do not have absolute meaning but only relative meaning:

part of $\psi(x)$	relative amplitude	normalized amplitude
incoming, $\mathrm{e}^{\mathrm{i}kx}$, $x < 0$	$1 + \mathrm{i}\tan(\alpha)$	1
transmitted, $\mathrm{e}^{\mathrm{i}kx}$, $x > 0$	1	$\dfrac{1}{1 + \mathrm{i}\tan(\alpha)}$
reflected, $\mathrm{e}^{-\mathrm{i}kx}$, $x < 0$	$-\mathrm{i}\tan(\alpha)$	$\dfrac{-\mathrm{i}\tan(\alpha)}{1 + \mathrm{i}\tan(\alpha)}$

(5.4.51)

In the last column, we have the amplitudes that result from normalizing the incoming part to unit amplitude so that the transmitted and reflected amplitudes are then relative to the incoming one. They are

$$\begin{aligned} \text{transmission}: \quad & \frac{1}{1+\mathrm{i}\tan(\alpha)} = \mathrm{e}^{-\mathrm{i}\alpha}\cos(\alpha)\,, \\ \text{reflection}: \quad & \frac{-\mathrm{i}\tan(\alpha)}{1+\mathrm{i}\tan(\alpha)} = -\mathrm{i}\,\mathrm{e}^{-\mathrm{i}\alpha}\sin(\alpha)\,, \end{aligned} \tag{5.4.52}$$

and we find the respective probabilities by squaring,

$$\begin{aligned} \text{prob(transmission)} &= \left|\,\mathrm{e}^{-\mathrm{i}\alpha}\cos(\alpha)\right|^2 = \cos(\alpha)^2 = \frac{k^2}{\kappa^2+k^2}\,, \\ \text{prob(reflection)} &= \left|\,-\mathrm{i}\,\mathrm{e}^{-\mathrm{i}\alpha}\sin(\alpha)\right|^2 = \sin(\alpha)^2 = \frac{\kappa^2}{\kappa^2+k^2}\,. \end{aligned} \tag{5.4.53}$$

Note that the qualitative aspects are as could have been expected. If the potential is of large relative strength, $\kappa \gg k$, the reflection probability is very large; if it is relatively weak, $\kappa \ll k$, then reflection is unlikely and transmission highly probable. In the force-free limit, $\kappa = 0$, there is no reflection at all, of course.

Note also that the sign of κ has no bearing on these probabilities. For either sign in

$$H = \frac{1}{2M}P^2 \mp \frac{\hbar^2\kappa}{M}\delta(X)\,, \tag{5.4.54}$$

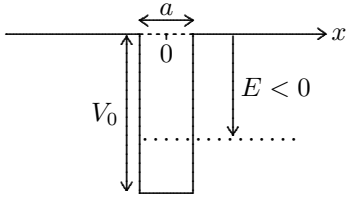
we get the same probabilities for reflection and transmission. This is not really a generic feature, but it is typical: Usually, attractive and repulsive potentials are difficult to distinguish on their *scattering* properties alone.

This last sentence introduces some terminology. The states for $E > 0$, where we have a *continuum* of energy eigenvalues, are called *scattering states* because they refer to physical situations of the kind just discussed: an incoming probability amplitude wave combined with transmitted and reflected ones. By contrast, the states to *discrete* energy eigenvalues, in this example only the one for $E = -(\hbar\kappa)^2/(2M)$, are *bound states*; their position wave functions decrease rapidly (usually exponentially) when you move away from the area where the potential energy $V(x)$ is relevant.

5.5 Square-well potential

5.5.1 *Bound states*

Harkening back to the discussion at (5.4.14), let us now take a look at the situation of an attractive square-well potential of finite width a and finite depth V_0:



$$V(x) = \begin{cases} 0 & \text{for} \quad |x| > \frac{1}{2}a\,, \\ -V_0 < 0 & \text{for} \quad |x| < \frac{1}{2}a\,. \end{cases} \tag{5.5.1}$$

We focus on the *bound states*, $E < 0$,

$$H\big|E\big\rangle = \big|E\big\rangle E \tag{5.5.2}$$

of the Hamilton operator

$$H = \frac{1}{2M}P^2 + V(X)\,. \tag{5.5.3}$$

Since $\langle H\rangle = E$, then, and $\big\langle P^2\big\rangle > 0$ while $0 > \langle V\rangle > -V_0$, we have

$$0 > E > -V_0 \tag{5.5.4}$$

or $E < 0$ and $E + V_0 > 0$.

The Schrödinger eigenvalue equation for $\psi_E(x) = \langle x|E\rangle$ is

$$\left(-\frac{\hbar^2}{2M}\frac{\partial^2}{\partial x^2} + V(x)\right)\psi_E(x) = E\psi_E(x)\,, \tag{5.5.5}$$

that is,

$$\begin{aligned} \frac{\partial^2}{\partial x^2}\psi_E(x) &= -\frac{2ME}{\hbar^2}\psi_E(x) && \text{for} \quad |x| > \frac{1}{2}a\,, \\ \frac{\partial^2}{\partial x^2}\psi_E(x) &= -\frac{2M(V_0+E)}{\hbar^2}\psi_E(x) && \text{for} \quad |x| < \frac{1}{2}a\,. \end{aligned} \tag{5.5.6}$$

Upon introducing the energy parameters κ and k by means of

$$\begin{aligned} E &= -\frac{(\hbar\kappa)^2}{2M} \quad \text{with } \kappa > 0\,, \\ E + V_0 &= +\frac{(\hbar k)^2}{2M} \quad \text{with } k > 0\,, \end{aligned} \tag{5.5.7}$$

we have

$$\frac{\partial^2}{\partial x^2}\psi_E(x) = \begin{cases} +\kappa^2\psi_E(x) & \text{for} \quad |x| > \frac{1}{2}a\,, \\ -k^2\psi_E(x) & \text{for} \quad |x| < \frac{1}{2}a\,. \end{cases} \tag{5.5.8}$$

Therefore, the even solutions are of the form

$$\psi_{\text{even}}(x) = \begin{cases} A\,\mathrm{e}^{-\kappa|x|} & \text{for} \quad |x| > \frac{1}{2}a\,, \\ B\cos(kx) & \text{for} \quad |x| < \frac{1}{2}a\,, \end{cases} \tag{5.5.9}$$

and the odd solutions are of the form

$$\psi_{\text{odd}}(x) = \begin{cases} A\,\mathrm{sgn}(x)\,\mathrm{e}^{-\kappa|x|} & \text{for} \quad |x| > \frac{1}{2}a\,, \\ B\sin(kx) & \text{for} \quad |x| < \frac{1}{2}a\,, \end{cases} \tag{5.5.10}$$

where A and B are to be chosen such that the respective $\psi(x)$s are continuous and also their derivatives,

$$\begin{aligned} \frac{\partial}{\partial x}\psi_{\text{even}}(\kappa) &= \begin{cases} -A\kappa\,\mathrm{sgn}(x)\,\mathrm{e}^{-\kappa|x|} & \text{for} \quad |x| > \frac{1}{2}a\,, \\ -Bk\sin(kx) & \text{for} \quad |x| < \frac{1}{2}a\,, \end{cases} \\ \frac{\partial}{\partial x}\psi_{\text{odd}}(x) &= \begin{cases} -A\kappa\,\mathrm{e}^{-\kappa|x|} & \text{for} \quad |x| > \frac{1}{2}a\,, \\ Bk\cos(kx) & \text{for} \quad |x| < \frac{1}{2}a\,. \end{cases} \end{aligned} \tag{5.5.11}$$

The continuity of $\psi(x)$ at $x = \frac{1}{2}a$ requires

$$\begin{aligned} \text{even}: \quad & A\,\mathrm{e}^{-\frac{1}{2}\kappa a} = B\cos\bigl(\tfrac{1}{2}ka\bigr)\,, \\ \text{odd}: \quad & A\,\mathrm{e}^{-\frac{1}{2}\kappa a} = B\sin\bigl(\tfrac{1}{2}ka\bigr)\,, \end{aligned} \tag{5.5.12}$$

and the continuity of $\partial\psi/\partial x$ at $x = \frac{1}{2}a$ requires

$$\begin{aligned} \text{even}: \quad & -\kappa A\,\mathrm{e}^{-\frac{1}{2}\kappa a} = -kB\sin\bigl(\tfrac{1}{2}ka\bigr)\,, \\ \text{odd}: \quad & -\kappa A\,\mathrm{e}^{-\frac{1}{2}\kappa a} = kB\cos\bigl(\tfrac{1}{2}ka\bigr)\,. \end{aligned} \tag{5.5.13}$$

Together,

$$\text{even:} \quad \kappa = k\tan\bigl(\tfrac{1}{2}ka\bigr)\,, \qquad \text{odd:} \quad \kappa = -k\cot\bigl(\tfrac{1}{2}ka\bigr)\,. \tag{5.5.14}$$

It is convenient to write $\vartheta = \frac{1}{2}ka$ and introduce θ in

$$\Bigl(\frac{1}{2}\kappa a\Bigr)^2 = -\frac{2ME}{\hbar^2}\Bigl(\frac{a}{2}\Bigr)^2 = \underbrace{\frac{2MV_0}{\hbar^2}\Bigl(\frac{a}{2}\Bigr)^2}_{=\theta^2} - \underbrace{\frac{2M(E+V_0)}{\hbar^2}\Bigl(\frac{a}{2}\Bigr)^2}_{=\left(\frac{1}{2}ka\right)^2 = \vartheta^2} = \theta^2 - \vartheta^2\,. \tag{5.5.15}$$

Thus, we now parameterize the looked-for energy eigenvalues by parameter ϑ in accordance with

$$E = -V_0 + \frac{(\hbar k)^2}{2M} = -V_0 + \frac{(2\hbar/a)^2}{2M}\vartheta^2\,, \quad \vartheta > 0 \tag{5.5.16}$$

and summarize the essential details of the potential energy in the corresponding parameter θ defined by

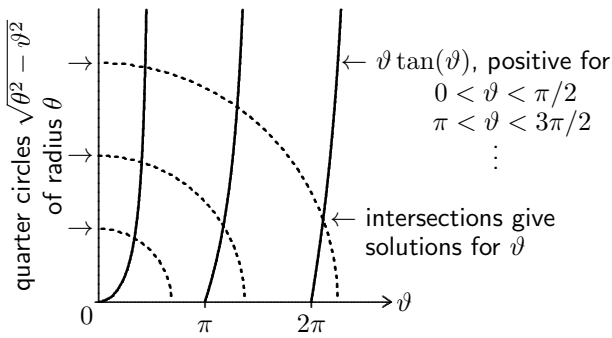
$$V_0 = \frac{(2\hbar/a)^2}{2M}\theta^2\,, \quad \theta > 0\,. \tag{5.5.17}$$

Equations (5.5.14) then turn into

$$\begin{aligned} \text{even}: &\quad \sqrt{\theta^2 - \vartheta^2} = \vartheta\tan(\vartheta)\,, \\ \text{odd}: &\quad \sqrt{\theta^2 - \vartheta^2} = -\vartheta\cot(\vartheta)\,. \end{aligned} \tag{5.5.18}$$

Their solutions, for ϑ, tell us the negative energy eigenvalues of the Hamilton operator in (5.5.3), that is, its *bound-state energies*.

These are transcendental equations that we cannot solve analytically, but it is easy to get solutions numerically. We can, however, establish a number of facts without knowing all these numerical details. For this purpose, we use a graphical representation of the two sides of the equation. For the *even* case, it is sketched as

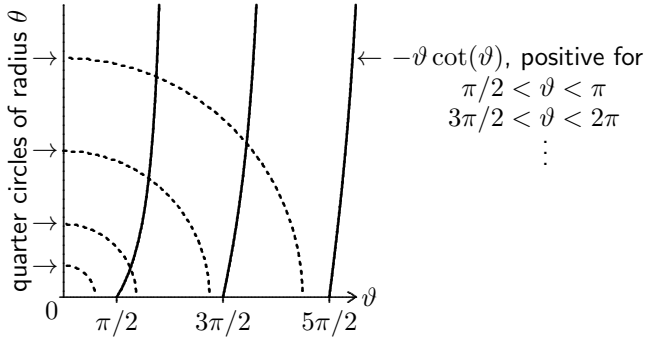


(5.5.19)

and we see that we have

$$\begin{aligned} &1 \text{ solution for } \theta < \pi\,, \\ &2 \text{ solutions for } \pi < \theta < 2\pi\,, \\ &3 \text{ solutions for } 2\pi < \theta < 3\pi\,, \\ &\qquad\vdots \\ &n \text{ solutions for } (n-1)\pi < \theta < n\pi\,. \end{aligned} \tag{5.5.20}$$

For *odd* wave functions, we have this picture



(5.5.21)

and here we read off that there is

$$\begin{aligned} &\text{no solution for } \theta < \pi/2\,,\\ &\text{1 solution for } \pi/2 < \theta < 3\pi/2\,,\\ &\text{2 solutions for } 3\pi/2 < \theta < 5\pi/2\,,\\ &\qquad\vdots\\ &n \text{ solutions for } (2n-1)\pi/2 < \theta < (2n+1)\pi/2\,. \end{aligned} \tag{5.5.22}$$

We get ϑ values for even solutions in the intervals $0 \cdots \frac{\pi}{2}$, $\pi \cdots \frac{3\pi}{2}$, $2\pi \cdots \frac{5\pi}{2}$, ... and ϑ values for odd solutions in the intervals $\frac{\pi}{2} \cdots \pi$, $\frac{3\pi}{2} \cdots 2\pi$, $\frac{5\pi}{2} \cdots 3\pi$, ... Accordingly, more and more bound states become available as θ increases, beginning with just one even bound state, then adding a first odd one, then a second even one, then a second odd one, and so forth. Altogether, there are n bound states if

$$n - 1 < \frac{2}{\pi}\theta < n \tag{5.5.23}$$

with either the same number of even and odd states (if $n = 2, 4, 6, \ldots$) or one more even state than there are odd states (if $n = 1, 3, 5, \ldots$).

Many of the details of this little calculation about a finite square-well potential are naturally pertinent to this potential only, such as the transcendental equations (5.5.18) for the values of ϑ and thus the precise energy eigenvalues. Other features are, however, generic. Any even potential energy function, $V(x) = V(-x)$, that supports bound states (that is an attractive potential) will have a lowest-energy state, a ground state, with an even wave function (no nodes), possibly an odd first excited state (one

node), then perhaps a second, even excited state (two nodes), and so forth. Potential energy functions that are not even can also lead to bound states, with zero nodes for the ground state, one node for the first excited state, two nodes for the second excited state, ... In fact, it is customary to label the bound states by their number of nodes. In this example of the finite square well, we would then write

$$\begin{aligned} &\psi_0(x)\,,\ \psi_2(x)\,,\ \psi_4(x)\,,\ \ldots \quad \text{for the even solutions,} \\ &\psi_1(x)\,,\ \psi_3(x)\,,\ \psi_5(x)\,,\ \ldots \quad \text{for the odd solutions,} \end{aligned} \tag{5.5.24}$$

with ϑ_0, ϑ_2, ϑ_4, ... solving the "even" equation in (5.5.18) and ϑ_1, ϑ_3, ϑ_5, ... solving the "odd" one, whereby

$$\frac{\pi}{2}k < \vartheta_k < \frac{\pi}{2}(k+1) \tag{5.5.25}$$

and $-V_0 < E_0 < E_1 < \cdots < 0$

$$\text{for} \quad E_k = -V_0 + \frac{(2\hbar/a)^2}{2M}\vartheta_k^2\,. \tag{5.5.26}$$

5.5.2 *Delta potential as a limit*

Let us briefly consider the limit of the delta function potential, that is,

$$V_0 \to \infty\,, \quad a \to 0 \quad \text{such that} \quad V_0 a \to \frac{\hbar^2\kappa}{M}\,, \tag{5.5.27}$$

where this κ is the one of Section 5.4. We write

$$V_0 a = \frac{\hbar^2\kappa}{M} \tag{5.5.28}$$

and then

$$\theta^2 = 2M\Big(\frac{a}{2\hbar}\Big)^2 V_0 = \frac{1}{2}\kappa a \tag{5.5.29}$$

leads to

$$\vartheta^2 + \big(\vartheta\tan(\vartheta)\big)^2 = \theta^2 = \frac{1}{2}\kappa a\,. \tag{5.5.30}$$

The term $\big(\vartheta\tan(\vartheta)\big)^2$ is $\propto \vartheta^4$ for small ϑ^2 so that

$$\vartheta^2 = \frac{1}{2}\kappa a - \left(\frac{1}{2}\kappa a\right)^2 + \cdots \tag{5.5.31}$$

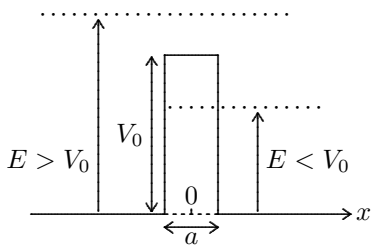
for small values of κa, as they are relevant here, in the limit of $a \to 0$ with κ fixed. This ϑ is, of course, the ϑ_0 of above, and the ground-state energy is

$$\begin{aligned} E_0 &= -V_0 + \frac{(2\hbar/a)^2}{2M}\vartheta_0^2 \\ &= -V_0 + \frac{(2\hbar/a)^2}{2M}\left(\frac{1}{2}\kappa a - \left(\frac{1}{2}\kappa a\right)^2 + \cdots\right) \\ &= \underbrace{-V_0 + \frac{\hbar^2\kappa}{Ma}}_{=0} - \frac{(\hbar\kappa)^2}{2M} + \underbrace{\cdots}_{\to 0 \text{ as } a \to 0} \\ &\to -\frac{(\hbar\kappa)^2}{2M} \quad \text{for} \quad a \to 0\,, \end{aligned} \tag{5.5.32}$$

as it should be. Indeed, we recover the earlier result for the ground-state energy of the attractive delta potential, the only negative energy eigenvalue available in this limit.

5.5.3 *Scattering states and tunneling*

As we did for the delta potential in Section 5.4.2, we could consider the scattering states for the attractive square-well potential of (5.5.1). Somewhat more interesting is, however, the repulsive square-well potential:



$$V(x) = \begin{cases} 0 & \text{for} \quad |x| > \frac{1}{2}a\,, \\ V_0 > 0 & \text{for} \quad |x| < \frac{1}{2}a\,, \end{cases} \tag{5.5.33}$$

where we expect a qualitative difference between high energies, $E > V_0$, and low energies, $0 < E < V_0$, because the region $-\frac{1}{2}a < x < \frac{1}{2}a$ could not be accessed in classical mechanics for low energies.

Indeed, classical intuition would let us expect that the atom is surely transmitted for $E > V_0$ and surely reflected for $E < V_0$. But what does quantum mechanics say about the actual situation?

The answer is given by the solution of the time-independent Schrödinger equation (5.5.5) for $V(x)$ of (5.5.33). In analogy with (5.5.7), we introduce

parameters k and κ by means of

$$E = \frac{(\hbar k)^2}{2M} \quad \text{with } k > 0\,,$$
$$\big|E - V_0\big| = \frac{(\hbar\kappa)^2}{2M} \quad \text{with } \kappa > 0\,. \tag{5.5.34}$$

This turns (5.5.5) into

$$\frac{\partial^2}{\partial x^2}\psi_E(x) = -k^2\psi_E(x) \quad \text{for} \quad |x| > \frac{1}{2}a \tag{5.5.35}$$

and

$$\frac{\partial^2}{\partial x^2}\psi_E(x) = \begin{cases} +\kappa^2\psi_E(x) & \text{for} \quad |x| < \frac{1}{2}a \text{ and } E < V_0\,, \\ -\kappa^2\psi_E(x) & \text{for} \quad |x| < \frac{1}{2}a \text{ and } E > V_0\,. \end{cases} \tag{5.5.36}$$

Guided by the discussion in Section 5.4.2, we are looking for a solution with the structure of the wave function in (5.4.49), depicted in (5.4.50): an incoming amplitude from the left, a transmitted amplitude outgoing to the right, and a reflected amplitude outgoing to the left. This is to say that we use

$$\psi_E(x) = \begin{cases} \mathrm{e}^{\mathrm{i}kx} + r\,\mathrm{e}^{-\mathrm{i}kx} & \text{for} \quad x < -\frac{1}{2}a\,, \\ t\,\mathrm{e}^{\mathrm{i}kx} & \text{for} \quad x > \frac{1}{2}a\,, \end{cases} \tag{5.5.37}$$

for the two regions where (5.5.35) applies. We thus normalize the incoming amplitude, that is, $\mathrm{e}^{\mathrm{i}kx}$ for $x < -\frac{1}{2}a$, to unity so that the *reflected amplitude* r and the *transmitted amplitude* t give us the respective probabilities by squaring,

$$\text{prob(reflection)} = |r|^2\,,$$
$$\text{prob(transmission)} = |t|^2\,. \tag{5.5.38}$$

In the region $-\frac{1}{2}a < x < \frac{1}{2}a$, the distinction between low and high energies is crucial, and in view of (5.5.36), we have

$$\psi_E(x) = \begin{cases} A\cosh(\kappa x) + B\sinh(\kappa x) & \text{for} \quad |x| < \frac{1}{2}a \text{ and } E < V_0\,, \\ A\cos(\kappa x) + B\sin(\kappa x) & \text{for} \quad |x| < \frac{1}{2}a \text{ and } E > V_0\,. \end{cases} \tag{5.5.39}$$

The four amplitude coefficients — r and t in (5.5.37) as well as A and B in (5.5.39) — are determined by the continuity of $\psi_E(x)$ and $\frac{\partial}{\partial x}\psi_E(x)$ at $x = -\frac{1}{2}a$ and $x = +\frac{1}{2}a$, which together are four conditions indeed.

The two continuity conditions at $x = -\frac{1}{2}a$ are

$$\begin{aligned}\psi_E\big(-\tfrac{1}{2}a\big) &= \mathrm{e}^{-\mathrm{i}\frac{1}{2}ka} + r\,\mathrm{e}^{\mathrm{i}\frac{1}{2}ka} \\ &= \begin{cases} A\cosh\big(\tfrac{1}{2}\kappa a\big) - B\sinh\big(\tfrac{1}{2}\kappa a\big) & \text{for} \quad E < V_0\,, \\ A\cos\big(\tfrac{1}{2}\kappa a\big) - B\sin\big(\tfrac{1}{2}\kappa a\big) & \text{for} \quad E > V_0 \end{cases}\end{aligned} \tag{5.5.40}$$

and

$$\begin{aligned}\frac{1}{\kappa}\frac{\partial\psi_E}{\partial x}\big(-\tfrac{1}{2}a\big) &= \frac{\mathrm{i}k}{\kappa}\Big(\mathrm{e}^{-\mathrm{i}\frac{1}{2}ka} - r\,\mathrm{e}^{\mathrm{i}\frac{1}{2}ka}\Big) \\ &= \begin{cases} -A\sinh\big(\tfrac{1}{2}\kappa a\big) + B\cosh\big(\tfrac{1}{2}\kappa a\big) & \text{for} \quad E < V_0\,, \\ A\sin\big(\tfrac{1}{2}\kappa a\big) + B\cos\big(\tfrac{1}{2}\kappa a\big) & \text{for} \quad E > V_0\,. \end{cases}\end{aligned} \tag{5.5.41}$$

And at $x = +\frac{1}{2}a$, we have the continuity conditions

$$\begin{aligned}\psi_E\big(\tfrac{1}{2}a\big) &= t\,\mathrm{e}^{\mathrm{i}\frac{1}{2}ka} \\ &= \begin{cases} A\cosh\big(\tfrac{1}{2}\kappa a\big) + B\sinh\big(\tfrac{1}{2}\kappa a\big) & \text{for} \quad E < V_0\,, \\ A\cos\big(\tfrac{1}{2}\kappa a\big) + B\sin\big(\tfrac{1}{2}\kappa a\big) & \text{for} \quad E > V_0 \end{cases}\end{aligned} \tag{5.5.42}$$

and

$$\begin{aligned}\frac{1}{\kappa}\frac{\partial\psi_E}{\partial x}\big(\tfrac{1}{2}a\big) &= \frac{\mathrm{i}k}{\kappa}\,t\,\mathrm{e}^{\mathrm{i}\frac{1}{2}ka} \\ &= \begin{cases} A\sinh\big(\tfrac{1}{2}\kappa a\big) + B\cosh\big(\tfrac{1}{2}\kappa a\big) & \text{for} \quad E < V_0\,, \\ -A\sin\big(\tfrac{1}{2}\kappa a\big) + B\cos\big(\tfrac{1}{2}\kappa a\big) & \text{for} \quad E > V_0\,. \end{cases}\end{aligned} \tag{5.5.43}$$

We leave the high-energy case of $E > V_0$ to Exercise 111 and here take a look at the low-energy case of $E < V_0$, for which (5.5.40) and (5.5.41) are compactly summarized by

$$\mathrm{e}^{\mathrm{i}\frac{1}{2}ka}\begin{pmatrix} 1 & 1 \\ \dfrac{\mathrm{i}k}{\kappa} & -\dfrac{\mathrm{i}k}{\kappa} \end{pmatrix}\begin{pmatrix} \mathrm{e}^{-\mathrm{i}ka} \\ r \end{pmatrix} = \begin{pmatrix} \cosh\big(\tfrac{1}{2}\kappa a\big) & -\sinh\big(\tfrac{1}{2}\kappa a\big) \\ -\sinh\big(\tfrac{1}{2}\kappa a\big) & \cosh\big(\tfrac{1}{2}\kappa a\big) \end{pmatrix}\begin{pmatrix} A \\ B \end{pmatrix} \tag{5.5.44}$$

and (5.5.42) and (5.5.43) by

$$\mathrm{e}^{\mathrm{i}\frac{1}{2}ka}\, t \begin{pmatrix} 1 \\ \dfrac{\mathrm{i}k}{\kappa} \end{pmatrix} = \begin{pmatrix} \cosh\bigl(\frac{1}{2}\kappa a\bigr) & \sinh\bigl(\frac{1}{2}\kappa a\bigr) \\ \sinh\bigl(\frac{1}{2}\kappa a\bigr) & \cosh\bigl(\frac{1}{2}\kappa a\bigr) \end{pmatrix} \begin{pmatrix} A \\ B \end{pmatrix}. \tag{5.5.45}$$

Since the two 2×2 matrices on the right are inverses of each other, we can immediately solve (5.5.44) for A and B and insert them into (5.5.45). The outcome

$$t \begin{pmatrix} 1 \\ \dfrac{\mathrm{i}k}{\kappa} \end{pmatrix} = \begin{pmatrix} \cosh(\kappa a) & \sinh(\kappa a) \\ \sinh(\kappa a) & \cosh(\kappa a) \end{pmatrix} \begin{pmatrix} 1 & 1 \\ \dfrac{\mathrm{i}k}{\kappa} & -\dfrac{\mathrm{i}k}{\kappa} \end{pmatrix} \begin{pmatrix} \mathrm{e}^{-\mathrm{i}ka} \\ r \end{pmatrix} \tag{5.5.46}$$

states two ways of expressing the transmitted amplitude t in terms of the reflected amplitude r,

$$\begin{aligned} t &= \bigl(\mathrm{e}^{-\mathrm{i}ka} + r\bigr)\cosh(\kappa a) + \frac{\mathrm{i}k}{\kappa}\bigl(\mathrm{e}^{-\mathrm{i}ka} - r\bigr)\sinh(\kappa a) \\ &= \bigl(\mathrm{e}^{-\mathrm{i}ka} - r\bigr)\cosh(\kappa a) + \frac{\kappa}{\mathrm{i}k}\bigl(\mathrm{e}^{-\mathrm{i}ka} + r\bigr)\sinh(\kappa a)\,, \end{aligned} \tag{5.5.47}$$

implying first

$$r = \frac{-\mathrm{i}(\kappa^2 + k^2)\sinh(\kappa a)\,\mathrm{e}^{-\mathrm{i}ka}}{2\kappa k\cosh(\kappa a) + \mathrm{i}(\kappa^2 - k^2)\sinh(\kappa a)} \tag{5.5.48}$$

and then

$$t = \frac{2\kappa k\,\mathrm{e}^{-\mathrm{i}ka}}{2\kappa k\cosh(\kappa a) + \mathrm{i}(\kappa^2 - k^2)\sinh(\kappa a)}\,. \tag{5.5.49}$$

If one wishes, one can now get the values of A and B, but we skip that.

As required by (5.5.38), we square r and t to get the probabilities for reflection and transmission,

$$\begin{aligned} \text{prob(reflection)} &= \left(1 + \left(\frac{\kappa^2 + k^2}{2\kappa k}\sinh(\kappa a)\right)^{-2}\right)^{-1}, \\ \text{prob(transmission)} &= \left(1 + \left(\frac{\kappa^2 + k^2}{2\kappa k}\sinh(\kappa a)\right)^{2}\right)^{-1}, \end{aligned} \tag{5.5.50}$$

which have unit sum, a basic check of consistency. Contrary to the classical expectation stated after (5.5.33), there is a nonzero transmission probability in the low-energy situation to which (5.5.50) applies and, accordingly,

the probability of reflection is less than 100%. This phenomenon, first noticed by Hund* in 1926/1927, is often referred to as "tunneling through the potential barrier" or simply known as the *tunnel effect*. More than just being an entertaining prediction of quantum theory, it is a real physical phenomenon that is exploited in devices, such as Esaki's† tunnel diode, the tunnel transistor, or Binnig's‡ and Rohrer's§ scanning tunneling microscope.

In this context, the transmission probability is known as the *tunneling probability*. It is rather small, unless the argument of the hyperbolic sine in (5.5.50) is small, but typically it is not. Consider, for example, the extreme situation of very low energy, $0 \lesssim E \ll V_0$, when $0 \lesssim k \ll \kappa \cong \sqrt{2MV_0/\hbar^2}$ and therefore

$$\text{prob(transmission)} \cong \left(\frac{2k}{\kappa \sinh(\kappa a)}\right)^2 = \frac{4E/V_0}{\sinh\bigl(\sqrt{2MV_0}\,a/\hbar\bigr)^2}\,. \qquad (5.5.51)$$

Owing to the squared hyperbolic sine in the denominator, this tunneling probability is exponentially small if the barrier width a is not tiny on the scale set by the height V_0 of the potential.

5.6 Stern–Gerlach experiment revisited

Finally, we return to the Stern–Gerlach experiment of (2.1.1) and treat it with the quantum formalism. We are content with an approximate description — a model — that captures the main features. While we use quantum dynamics for the transverse z direction (position operator Z and momentum operator P), we regard the longitudinal motion in the y direction as classical, $y = y(t)$, and we ignore the x direction altogether. The magnetic field $\boldsymbol{B}(\boldsymbol{r})$ is then reduced to its z component, which depends on the transverse coordinate z and acquires a time dependence through $y(t)$,

$$\boldsymbol{B} \to \boldsymbol{e}_z\, B_z\bigl(x = 0, y(t), z\bigr)\,. \qquad (5.6.1)$$

In addition, we assume that in the region explored by the atom, the z component has a z-independent gradient,

$$B_z\bigl(x = 0, y(t), z\bigr) \cong B_0(t) + b(t)z\,, \qquad (5.6.2)$$

*Friedrich Hermann Hund (1896–1997) †Leo Esaki (b. 1925)
‡Gerd Binnig (b. 1947) §Heinrich Rohrer (1933–2013)

where $B_0(t)$ is the on-axis field strength at $z = 0$ and $b(t)$ is the gradient. Therefore, the potential energy of the magnetic dipole of the atom, $\boldsymbol{\mu} = \mu\boldsymbol{\sigma}$, in the magnetic field is approximated by

$$-\boldsymbol{\mu} \cdot \boldsymbol{B} \cong -\mu\big(B_0(t) + b(t)z\big)\sigma_z \,, \tag{5.6.3}$$

and then

$$H = \frac{1}{2M}P^2 - \mu\big(B_0(t) + b(t)Z\big)\sigma_z \tag{5.6.4}$$

is the Hamilton operator that models the atom traversing the Stern–Gerlach apparatus.

The Heisenberg equations of motion obeyed by $Z(t)$, $P(t)$, and $\sigma_z(t)$ are

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}Z(t) &= \frac{1}{\mathrm{i}\hbar}\big[Z(t), H\big] = \frac{1}{M}P(t) \,, \\ \frac{\mathrm{d}}{\mathrm{d}t}P(t) &= \frac{1}{\mathrm{i}\hbar}\big[P(t), H\big] = \mu b(t)\sigma_z(t) \,, \\ \frac{\mathrm{d}}{\mathrm{d}t}\sigma_z(t) &= \frac{1}{\mathrm{i}\hbar}\big[\sigma_z(t), H\big] = 0 \,. \end{aligned} \tag{5.6.5}$$

Since $\sigma_z(t)$ is actually constant in time,

$$\sigma_z(t) = \sigma_z(0) \,, \tag{5.6.6}$$

we find $P(t)$ by integration,

$$P(t) = P(0) + \int_0^t \mathrm{d}t' \, \mu b(t')\sigma_z(0) \,, \tag{5.6.7}$$

and another integration yields

$$\begin{aligned} Z(t) &= Z(0) + \int_0^t \mathrm{d}t' \, \frac{1}{M}P(t') \\ &= Z(0) + \frac{t}{M}P(0) + \frac{1}{M}\int_0^t \mathrm{d}t' \int_0^{t'} \mathrm{d}t'' \, \mu b(t'')\sigma_z(0) \,. \end{aligned} \tag{5.6.8}$$

Here, the initial time $t = 0$ is before the atom enters the apparatus, and $b(t)$ is nonzero only while the atom traverses the apparatus. In the double integral, t' is always later than t'' so that

$$\int_0^t \mathrm{d}t' \int_0^{t'} \mathrm{d}t'' \, b(t'') = \int_0^t \mathrm{d}t'' \int_{t''}^t \mathrm{d}t' \, b(t'') = \int_0^t \mathrm{d}t'' \, (t - t'')b(t'') \,. \tag{5.6.9}$$

We recognize that

$$\Delta p(t) = \int_0^t \mathrm{d}t'\, \mu b(t') \tag{5.6.10}$$

is the momentum transferred by the force $\mu b(t')$ between $t' = 0$ and $t' = t$, and

$$\Delta z(t) = \frac{1}{M} \int_0^t \mathrm{d}t'\, (t - t') \mu b(t') \tag{5.6.11}$$

is the displacement associated with this momentum transfer in

$$\begin{aligned} P(t) &= P(0) + \Delta p(t)\, \sigma_z(0)\,, \\ Z(t) &= Z(0) + \frac{t}{M} P(0) + \Delta z(t)\, \sigma_z(0)\,. \end{aligned} \tag{5.6.12}$$

When the atom has the spin up in the z direction, $\sigma_z = +1$, the momentum is changed by $+\Delta p(t)$, and by $-\Delta p(t)$ when the spin is down, $\sigma_z = -1$; the respective displacements are $+\Delta z(t)$ and $-\Delta z(t)$.

An atom that has its spin initially aligned in another direction, perhaps the x direction, arrives in a superposition of up and down in the z direction — recall that $|{\uparrow_x}\rangle = \big(|{\uparrow_z}\rangle + |{\downarrow_z}\rangle\big)/\sqrt{2}$ — and then the atom's center-of-mass motion (variables Z and P) gets entangled with its spin degree of freedom (variables σ_x, σ_y, σ_z). As a consequence of this entanglement, we can infer the z component of the atom's spin from a measurement of its momentum or position. This is the essence of the Stern–Gerlach experiment.

The experiment succeeds provided that the accumulated momentum transfer $\Delta p(T)$, at a late time $t = T$ after the atom has exited from the apparatus, is substantially larger than the momentum spread of the arriving atoms, which is estimated in Exercise 93. Then the two partial beams separate as indicated in (2.1.14). It follows that a successful Stern–Gerlach experiment requires a field gradient that is strong enough to ensure the beam separation. This condition can be met, and Stern and Gerlach did meet it in 1922.

This page intentionally left blank

Exercises with Hints

Chapter 1

1 Answer the question on page 4: How likely are four T's and six R's in the next sequence of ten?

2 If T clicks and R clicks occur with equal probability, what is the probability of observing exactly n clicks of each kind in a sequence of $2n$ clicks?

Chapter 2

3 Show that the commutativity observed in (2.4.25) is generally true. That is, if X, Y, Z are $n \times n$ square matrices such that

$$XY = 1 \quad \text{and} \quad YZ = 1\,,$$

then it follows (how?) that $X = Z$, and we write $X = Z = Y^{-1}$.

4 Compare the following two situations:

(A) A beam of atoms, half of which are preselected as "+ in z" and the other half as "− in z," is sent through a Stern–Gerlach apparatus that
  (i) sorts in the z direction,
  (ii) sorts in the x direction.

(B) A beam of atoms, all of which are preselected as "+ in x," is sent through a Stern–Gerlach apparatus that
  (i) sorts in the z direction,
  (ii) sorts in the x direction.

Do the measurement results enable you to tell apart situations (A) and (B)?

5 As a sequel to (2.5.16), show that one also gets 50%:50% probability when atoms preselected as "+ in x" or "− in x" are sent through a y-sorting Stern–Gerlach apparatus.

6 How would you use a homogeneous magnetic field in the x direction combined with a z-deflecting Stern–Gerlach magnet to carry out that $\pm y$ sorting?

7 Exploit the normalization and orthogonality conditions of (2.5.30)–(2.5.32) to show that the statement is generally correct.

8 Illustrate (2.5.39) by writing out explicitly the six projection matrices for the "± in x," "± in y," and "± in z" columns and adding each pair.

9 Calculate the probability p_n in (2.6.6) for $n = 4, 8, 16$, and 32.

10 Compare the approximation for p_n in (2.6.7) with the exact numbers of Exercise 9. Is the error of the size you expect?

11 Atoms that have been preselected as "+ in z" are successively passed through first a "+ in x" selector and then a "− in z" selector. Which fraction of the atoms is let through?

12 For the state described by the ket in (2.7.6) with $\alpha = \frac{3}{5}$ and $\beta = \frac{4}{5}$, determine the probability amplitudes, and the resulting probabilities, of finding $\uparrow_x$ and $\downarrow_x$, or $\uparrow_y$ and $\downarrow_y$, or $\uparrow_z$ and $\downarrow_z$. Then, confirm that the paired probabilities have unit sum.

13 Show that (2.8.19) must hold for any bra $\langle 1|$ and ket $|2\rangle$ for consistency with (2.8.13) and (2.8.14).

14 Repeat (2.8.20) and (2.8.21) for the y kets and bras.

15 Consider n pairs of kets, the kth pair denoted by $|a_k\rangle$ and $|b_k\rangle$, that are jointly defined by

$$\big|a_k\big\rangle = \big|{\uparrow_z}\big\rangle u_k^* + \big|{\downarrow_z}\big\rangle v_k\,, \qquad \big|b_k\big\rangle = \big|{\uparrow_z}\big\rangle v_k^* - \big|{\downarrow_z}\big\rangle u_k\,,$$

for $k = 1, 2, \ldots, n$, where the amplitudes u_k and v_k are complex numbers that are constrained by $|u_k|^2 + |v_k|^2 = 1$ but are arbitrary otherwise. How

large are the probabilities $\big|\langle a_k|b_k\rangle\big|^2$? How are the probability amplitudes $\langle a_j|a_k\rangle$ and $\langle b_j|b_k\rangle$ related to each other?

16 Many nonstandard matrix representations supplement the standard representation of (2.9.10), among them

$$\sigma_a \mathrel{\widehat{=}} \begin{pmatrix} \cos(\vartheta) & \mathrm{e}^{-\mathrm{i}\varphi_a}\sin(\vartheta) \\ \mathrm{e}^{\mathrm{i}\varphi_a}\sin(\vartheta) & -\cos(\vartheta) \end{pmatrix} \quad \text{for} \quad a = x, y, z\,,$$

which is more symmetrical than the standard representation inasmuch as the three matrices for σ_x, σ_y, and σ_z have the same structure and differ only by the respective phase factors $\mathrm{e}^{\mathrm{i}\varphi_x}$, $\mathrm{e}^{\mathrm{i}\varphi_y}$, and $\mathrm{e}^{\mathrm{i}\varphi_z}$. Find the common value of ϑ (with $0 < 2\vartheta < \pi$) and determine a consistent choice for φ_x, φ_y, and φ_z.

17 Show that $(\boldsymbol{a}\cdot\boldsymbol{\sigma})^2 = \boldsymbol{a}\cdot\boldsymbol{a} = \boldsymbol{a}^2$ for any numerical 3-vector $\boldsymbol{a}$, real or complex. What do you get for $\boldsymbol{a} = \boldsymbol{e}_1 + \mathrm{i}\boldsymbol{e}_2$?

18 Show that $\boldsymbol{a}\cdot\boldsymbol{\sigma}\ \ \boldsymbol{b}\cdot\boldsymbol{\sigma} = -\boldsymbol{b}\cdot\boldsymbol{\sigma}\ \ \boldsymbol{a}\cdot\boldsymbol{\sigma}$ if $\boldsymbol{a}$ and $\boldsymbol{b}$ are perpendicular to each other.

19 Express $(1 + \mathrm{i}\sigma_x)\sigma_y(1 - \mathrm{i}\sigma_x)$ as a linear function of $\boldsymbol{\sigma}$.

20 Same for $\frac{1}{4}(1 + \sigma_x)(1 + \sigma_y)(1 + \sigma_x)$. Square the outcome.

21 What is $\sigma_x\sigma_y\sigma_z$?

22 If $\boldsymbol{n}$ is a unit 3-vector, what is $\big(\frac{1}{2}(1 + \boldsymbol{n}\cdot\boldsymbol{\sigma})\big)^2$?

23 Show that

$$f(\sigma_x)\sigma_z = \sigma_z f(-\sigma_x)$$

holds for the products of the Pauli operator σ_z with *any* function of the Pauli operator σ_x.

24 Express the operator

$$A = \big|{\uparrow_x}\big\rangle\big\langle{\uparrow_z}\big| + \big|{\downarrow_x}\big\rangle\big\langle{\downarrow_z}\big|$$

as a linear function of $\boldsymbol{\sigma}$. What is A^2?

25 Operator A has eigenvalues a_1 and a_2 with eigenbra $\langle a_1|$ and eigenket $|a_2\rangle$. Show that $\langle a_1|a_2\rangle = 0$ if $a_1 \neq a_2$.

26 If operator A has an eigenket $|a\rangle$ to eigenvalue a, then $A^\dagger$ has an eigenbra $\langle a^*|$ to eigenvalue a^*. Why? How is $\langle a^*|$ related to $|a\rangle$?

27 Verify (2.11.32)–(2.11.35).

28 Show that, as a consequence of (2.11.35), every operator is a linear combination of

$$\big|{\uparrow_{\boldsymbol{e}}}\big\rangle\big\langle{\uparrow_{\boldsymbol{e}}}\big|\,,\ \big|{\uparrow_{\boldsymbol{e}}}\big\rangle\big\langle{\downarrow_{\boldsymbol{e}}}\big|\,,\ \big|{\downarrow_{\boldsymbol{e}}}\big\rangle\big\langle{\uparrow_{\boldsymbol{e}}}\big|\,,\ \big|{\downarrow_{\boldsymbol{e}}}\big\rangle\big\langle{\downarrow_{\boldsymbol{e}}}\big|\,.$$

That is, this quartet is a basis in the operator space.

29 For $\vartheta = \dfrac{\pi}{2}, \varphi = 0$ you have $\boldsymbol{e}\cdot\boldsymbol{\sigma} = \sigma_x$. Compare the kets $|{\uparrow_{\boldsymbol{e}}}\rangle$, $|{\downarrow_{\boldsymbol{e}}}\rangle$ in (2.11.26) and (2.11.31) with the kets $|{\uparrow_x}\rangle$, $|{\downarrow_x}\rangle$ found previously. Repeat for $\vartheta = \dfrac{\pi}{2}$, $\varphi = \dfrac{\pi}{2}$ when $\boldsymbol{e}\cdot\boldsymbol{\sigma} = \sigma_y$.

30 For the situation considered in Section 2.5, where the ket of (2.7.6) applies, show that

$$\langle\sigma_x\rangle = 2\,\mathrm{Re}(\alpha^*\beta)\ ,\quad \langle\sigma_y\rangle = 2\,\mathrm{Im}(\alpha^*\beta)\ ,\quad \langle\sigma_z\rangle = |\alpha|^2 - |\beta|^2\,.$$

What is their significance in the context of the measurement sketched in (2.5.42)?

31 Evaluate $\mathrm{tr}\big(|{\uparrow_x}\rangle\langle{\downarrow_y}|\big)$ and $\mathrm{tr}(1) = \mathrm{tr}\big(|{\uparrow_x}\rangle\langle{\uparrow_x}| + |{\downarrow_x}\rangle\langle{\downarrow_x}|\big)$.

32 Evaluate the trace of $X = 3 + \sigma_x + \sigma_z$ by using the matrix representations referring to σ_x, σ_y, and σ_z.

33 The operator Λ in (2.14.13) has eigenvalues λ_1 and λ_2. Show that their sum is the trace, $\mathrm{tr}(\Lambda) = \lambda_1 + \lambda_2$.

34 Which matrices represent $\sigma_x + \sigma_y + \sigma_z$ in the standard representation of (2.9.10) and the nonstandard representation of Exercise 16? Verify that the two matrices have the same trace and the same determinant. Use their values to find the eigenvalues of $\sigma_x + \sigma_y + \sigma_z$.

35 A source emits atoms such that each of them is either "+ in x" or "− in z," choosing randomly between these options, with equal chances for

both. What are the probabilities of (i) finding the next atom as "+ in x" or "− in x," (ii) finding the next atom as "+ in y" or "− in y," and (iii) finding the next atom as "+ in z" or "− in z," when the respective experiments are performed?

36 For ρ in (2.15.15), show that $\frac{1}{2} \leq \mathrm{tr}\big(\rho^2\big) \leq 1$. When does each "=" apply?

37 The statement in (2.15.15) applies to statistical operators. Show that, more generally, we have

$$A = a_0 + \boldsymbol{a} \cdot \boldsymbol{\sigma} \quad \text{with} \quad a_0 = \frac{1}{2}\mathrm{tr}(A) \quad \text{and} \quad \boldsymbol{a} = \frac{1}{2}\mathrm{tr}(\sigma A)$$

for any operator A so that the identity 1 and the three Pauli operators $\sigma_x, \sigma_y, \sigma_z$ are another basis is the operator space.

38 Express $f_1(\sigma_x) = \mathrm{e}^{\mathrm{i}\varphi\sigma_x}$ and $f_2(\sigma_x) = \dfrac{1 + \mathrm{i}\tau\sigma_x}{1 - \mathrm{i}\tau\sigma_x}$ with real parameters φ and τ, as linear functions of σ_x. How must φ and τ be related to each other in order to ensure $f_1(\sigma_x) = f_2(\sigma_x)$?

39 We define four hermitian operators $W_0, \ldots, W_3$ by

$$W_0 = \frac{1}{2}(1 + \sigma_x + \sigma_y + \sigma_z)\,,$$
$$W_1 = \sigma_x W_0 \sigma_x\,, \quad W_2 = \sigma_y W_0 \sigma_y\,, \quad W_3 = \sigma_z W_0 \sigma_z\,.$$

Show that they are normalized to unit trace, $\mathrm{tr}(W_j) = 1$ for $j = 0, 1, 2, 3$, and pairwise orthogonal in the sense of

$$\mathrm{tr}(W_j W_k) = 2\,\delta_{jk} \quad \text{for } j, k = 0, 1, 2, 3\,.$$

40 Show that these W_js are yet another basis in the operator space, that is, any operator Λ can be written as a weighted sum of the W_js,

$$\Lambda = \frac{1}{2}\sum_{j=0}^{3} \lambda_j W_j \quad \text{with } \lambda_j = \mathrm{tr}(\Lambda W_j)\,.$$

Express $\mathrm{tr}(\Lambda)$ and $\mathrm{tr}\big(\Lambda^\dagger \Lambda\big)$ in terms of the coefficients λ_j.

41 If the statistical operator is $\rho = \frac{1}{2}\sum_k r_k W_k$, what is the expectation value $\langle \Lambda \rangle$ in terms of the coefficients λ_j and r_k?

42 Atoms are prepared by blending $\frac{1}{2}$ of $\uparrow_x$ with $\frac{1}{3}$ of $\uparrow_y$ and $\frac{1}{6}$ of $\downarrow_z$. What is the resulting statistical operator?

43 For this preparation, what is the probability of "up" in direction $\boldsymbol{e} = \frac{1}{3}(2\,\boldsymbol{e}_1 - \boldsymbol{e}_2 + 2\,\boldsymbol{e}_3)$?

44 Show that the net effect of the nonselective z measurement in (2.17.1) can be described by

$$\rho_{\text{in}} \to \rho_{\text{out}} = \frac{1}{2}\big(\rho_{\text{in}} + \sigma_z \rho_{\text{in}} \sigma_z\big)\,.$$

How would you describe the effect of a nonselective x measurement?

45 Show the sufficiency of (2.18.17), that is, the ket in (2.18.8) is a product ket of the kind in (2.18.16) when (2.18.17) holds.

46 Express the two-atom ket in (2.18.18) in terms of up/down states for the x-direction, that is, as a sum of $|\uparrow_x\uparrow_x\rangle, |\uparrow_x\downarrow_x\rangle, \ldots$; repeat for the y-direction.

47 The two-party ket in (2.18.18) describes a maximally entangled state because the partial traces yield completely mixed states, $\rho^{(1)} = \text{tr}_2\big(|\,\rangle\langle\,|\big) = \frac{1}{2}$ and $\rho^{(2)} = \text{tr}_1\big(|\,\rangle\langle\,|\big) = \frac{1}{2}$. Show that $|\alpha\beta - \gamma\delta| = \frac{1}{2}$ whenever the ket in (2.18.8) refers to a maximally entangled state.

48 Verify that the kets in (2.20.8) are eigenkets of the top-row operators in Mermin's table (2.20.2).

49 How do you convert the kets in (2.20.8) into eigenkets for the other two rows or the columns in Mermin's table?

50 What are the eigenvalues, eigenkets, and eigenbras of $A^\dagger A$ and $AA^\dagger$ for A and $A^\dagger$ in (2.21.9) and (2.21.11), respectively? Show that $A^\dagger A = AA^\dagger$. Such operators are called *normal*; they represent, quite generally, physical properties.

51 Explain why $\text{tr}(A)$, the trace of operator A, is not a function of A in the sense of (2.21.17).

52 Guided by Exercise 28, state the operator basis that follows from the completeness relation in (2.21.4).

53 Consider the unitary operators in (2.22.3) and (2.22.6). If properties A and B are physically similar in the sense that $a_j = b_j$ for $j = 1, 2, \ldots, n$, then show that

$$BU_{ba} = U_{ba}A \quad \text{and} \quad AU_{ab} = U_{ab}B$$

hold.

54 Then, show that

$$U_{ba}f(A) = f(B)U_{ba} \quad \text{and} \quad f(A)U_{ab} = U_{ab}f(B)$$

under these circumstances.

55 Next, as an immediate extension of Exercise 54, show that

$$U^{-1}f(A)U = f(U^{-1}AU)$$

for any operator A representing a physical property and any unitary operator U.

56 The converse is also true: Show that

$$f(A) = \sum_{k=1}^{n} \big|a_k\big\rangle f(a_k) \big\langle a_k\big|$$

is unitary if $\big|f(a_k)\big|^2 = 1$ for $k = 1, 2, \ldots, n$.

57 An operator is known to be both unitary and hermitian. What can you say about its eigenvalues? Show that $A = \mathrm{e}^{\mathrm{i}\pi(1\,-\,A)/2}$ for every such hermitian-unitary operators A.

58 Given a normal operator A, that is $A^\dagger A = AA^\dagger$, see Exercise 50, show that there is a hermitian operator B such that A is a function of B.

59 If an operator is not normal, then either its eigenstates are not complete or its eigenstates are not pairwise orthogonal or both. Confirm that the two operators $\sigma_x + \mathrm{i}\sigma_y$ and $2\sigma_x + \mathrm{i}\sigma_y$ are not normal and check the properties of the eigenkets for both of them.

Chapter 3

60 Show that the Jacobi identity in (3.2.6) holds for all trios of operators X, Y, Z.

61 Now apply the Jacobi identity to $X = \boldsymbol{a} \cdot \boldsymbol{\sigma}$, $Y = \boldsymbol{b} \cdot \boldsymbol{\sigma}$, and $Z = \boldsymbol{c} \cdot \boldsymbol{\sigma}$ and derive an identity obeyed by the three numerical vectors $\boldsymbol{a}$, $\boldsymbol{b}$, and $\boldsymbol{c}$.

62 Show that

$$\mathrm{tr}\big([X,Y]Z\big) = \mathrm{tr}\big([Y,Z]X\big) = \mathrm{tr}\big([Z,X]Y\big)$$

for any trio X, Y, Z of operators. What is the common value of the three traces for the trio of Pauli operators in Exercise 61?

63 As a generalization of (3.3.1), consider a time-dependent shift of the origin on the energy scale,

$$H \to H + \hbar\Omega(t)\,.$$

How do you then relate $\langle \ldots, t|'$ to $\langle \ldots, t|$?

64 For a silver atom with magnetic moment $\boldsymbol{\mu}$, we have the magnetic interaction energy $-\boldsymbol{\mu} \cdot \boldsymbol{B}$ of (2.1.9). The relation between $\boldsymbol{\mu}$ and $\boldsymbol{\sigma}$ is

$$\boldsymbol{\mu} = -\mu_{\mathrm{B}}\boldsymbol{\sigma}\,,$$

where $\mu_{\mathrm{B}} = 9.274 \times 10^{-24}\,\mathrm{J/T}$ (Joule/Tesla) is the so-called *Bohr magneton*, and the negative sign comes about because electrons carry negative charge. How is $\boldsymbol{\omega}$ in (3.5.12) related to $\boldsymbol{B}$? How large is $\omega = |\boldsymbol{\omega}|$ for a magnetic field of $100\,\mathrm{G} = 0.01\,\mathrm{T}$?

65 For $H = \frac{1}{2}\hbar\omega\sigma_z$, state and solve the Schrödinger equations for $\langle \uparrow_x, t|$ and $\langle \downarrow_x, t|$.

66 For $\mathcal{H} = \hbar\omega \begin{pmatrix} 0 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & 0 \end{pmatrix}$ and $\psi(0) = \frac{1}{3}\begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix}$, find $\psi(t) = \begin{pmatrix} \alpha_1(t) \\ \alpha_2(t) \\ \alpha_3(t) \end{pmatrix}$.

67 For $\mathcal{H} = \hbar\omega \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$ and $\psi(0) = \frac{1}{\sqrt{3}}\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$, find $\psi(t)$.

68 For the 3×3 matrix $\mathcal{H}$ in Exercise 66, find the eigenvalues, eigencolumns, and eigenrows and use them to construct the unitary matrix $\mathrm{e}^{-\mathrm{i}\mathcal{H}t/\hbar}$ explicitly.

69 Apply your result to the initial column $\psi_0 = \frac{1}{3}\begin{pmatrix}1\\2\\2\end{pmatrix}$ and compare the outcome for $\psi(t)$ with what you found in Exercise 66 by solving the differential equation.

70 Do this also for the matrix $\mathcal{H}$ and the initial column in Exercise 67.

71 In (3.8.17), which transition frequency corresponds to an energy difference of 1 eV ?

72 In (3.9.34), we have the probability of finding $\uparrow_x\uparrow_x$ at time t if we have $\uparrow_x\uparrow_x$ at time $t=0$. In the same situation, what is the probability of finding $\uparrow_y\uparrow_y$ at time t if we have $\uparrow_y\uparrow_y$ at $t=0$?

73 Same question for $\uparrow_x\downarrow_x$ at $t=0$ and later.

Chapter 4

74 A model of the delta function is provided by $D(\phi) = \mathrm{e}^{-\frac{1}{2}\phi^2}$, which clearly has the properties required in (4.1.18)(a). Evaluate $\delta(x-x';\epsilon)$ for this case and discuss its properties. The gaussian integral

$$\int \mathrm{d}x\ \mathrm{e}^{-ax^2+2bx} = \sqrt{\pi/a}\ \mathrm{e}^{b^2/a} \quad \text{for} \quad \mathrm{Re}(a)\geq 0$$

will be useful.

75 While the graph in (4.1.22) is typical for a model of the delta function, there are models with rather different graphs. Consider

$$D(\phi) = \begin{cases} 1 \text{ for } |\phi|<1\,, \\ 0 \text{ for } |\phi|>1\,. \end{cases}$$

Find the corresponding $\delta(x-x';\epsilon)$ and graph it.

76 Find the constant in $\psi(x) = (\text{constant})\,\mathrm{e}^{-\kappa|x|}$, $\kappa>0$, that is needed for proper normalization of $\psi(x)$. Then, evaluate $\langle X\rangle$ and $\langle X^2\rangle$.

77 For the wave function of Exercise 76, evaluate $\left\langle \mathrm{e}^{\mathrm{i}kX} \right\rangle$ for real k and extract $\langle X^n \rangle$ for $n = 0, 1, 2, \ldots$ by expanding in powers of k.

78 Atoms in a gas at room temperature have an average kinetic energy $p^2/(2M)$ (M = mass of the atom) of about 0.025 eV. What would be the corresponding de Broglie wavelength for silver atoms? For helium atoms?

79 At (4.6.18), we argue that $\left|\psi(x)\right|^2 \to 0$ when $|x| \to \infty$ because of the normalization of $\left|\psi(x)\right|^2$ to unit integral — an argument that is, in fact, a bit sloppy. Show this by giving an example of a function $\psi(x)$, for which (4.6.19) holds, but $\left|\psi(x)\right|^2 \to 0$ for $x \to \pm\infty$ is not true. How, then, can you ensure the null result of (4.6.18)?

80 Show that the following expressions are equivalent:

$$\begin{aligned} \left\langle P^2 \right\rangle &= -\hbar^2 \int \mathrm{d}x\; \psi(x)^* \left(\frac{\partial}{\partial x}\right)^2 \psi(x) \\ &= -\hbar^2 \int \mathrm{d}x\; \psi(x) \left(\frac{\partial}{\partial x}\right)^2 \psi(x)^* = \hbar^2 \int \mathrm{d}x\; \left|\frac{\partial}{\partial x}\psi(x)\right|^2 . \end{aligned}$$

Should you be bothered by the lesson of Exercise 79?

81 For $\psi(x)$ in Exercise 76, find $\psi(p)$ and then calculate $\langle P \rangle$ and $\left\langle P^2 \right\rangle$ in at least two different ways each.

82 What is $\delta X\, \delta P$ for the state specified by the wave function $\psi(x)$ in Exercise 76?

83 In the derivation of the general uncertainty relation in (4.7.15), we divide by $\delta A\, \delta B$ at one point, which is only permissible if both spreads are truly positive. Why does this restriction not matter in the end? That is, show that the uncertainty relation applies also when $\delta A = 0$, or $\delta B = 0$, or $\delta A = 0$ and $\delta B = 0$.

84 Replace $A \to \lambda^{-1}A + \lambda B$ and $B \to \lambda^{-1}A - \lambda B$ in Robertson's inequality (4.7.15) with $\lambda > 0$. Find the value of λ, for which the resulting inequality is strictest, and so arrive at Schrödinger's uncertainty relation,

$$\delta A\, \delta B \geq \frac{1}{2}\sqrt{\left\langle \mathrm{i}[A,B] \right\rangle^2 + \Big(\big\langle (AB + BA) \big\rangle - 2\,\langle A \rangle \langle B \rangle\Big)^2} .$$

85 For two hermitian operators A, B and two complex coefficients α, β, generalize (5.1.65) to

$$(\alpha A + \beta B)\big|\ \big\rangle = \big|\ \big\rangle a + \big|{}_{\perp}\big\rangle b\,.$$

Demonstrate that $b = \langle{}_{\perp}|(\alpha A + \beta B)|\ \rangle$ and confirm that $a = \alpha\langle A\rangle + \beta\langle B\rangle$ and

$$|b|^2 = |\alpha|^2(\delta A)^2 + |\beta|^2(\delta B)^2 + \mathrm{Re}(\alpha^*\beta)\Big(\big\langle(AB+BA)\big\rangle - 2\,\langle A\rangle\langle B\rangle\Big) + \mathrm{Im}(\alpha^*\beta)\,\big\langle \mathrm{i}[A,B]\big\rangle\ ,$$

where $\langle A\rangle = \langle\ |A|\ \rangle$, for instance.

86 Now, recover Robertson's inequality from $|b|^2 \geq 0$ for $\mathrm{Re}(\alpha^*\beta) = 0$.

87 Next, consider the case of $\mathrm{Im}(\alpha^*\beta) = 0$ and infer that

$$|\alpha|^2(\delta A)^2 + |\beta|^2(\delta B)^2 \geq \Big|\alpha\beta\Big(\big\langle(AB+BA)\big\rangle - 2\,\langle A\rangle\langle B\rangle\Big)\Big|$$

and then show that

$$|\alpha|^2(\delta A)^2 + |\beta|^2(\delta B)^2 \geq \frac{1}{2}\Big|\langle{}_{\perp}|(\alpha A + \beta B)|\ \rangle\Big|^2 .$$

Uncertainty relations of this kind, for sums of variances rather than products, were first established by Maccone* and Pati.†

88 Further, demonstrate that $2\,\delta A\,\delta B \geq \big|\langle(AB+BA)\rangle - 2\langle A\rangle\langle B\rangle\big|$ and then conclude that

$$|\alpha|\delta A + |\beta|\delta B \geq \Big|\langle{}_{\perp}|(\alpha A + \beta B)|\ \rangle\Big|\,.$$

89 Explain why the inequalities of Maccone–Pati type in Exercises 87 and 88 remain valid when the particular $|{}_{\perp}\rangle$ of Exercise 85 is replaced by any $|{}_{\perp}{}'\rangle$ that is orthogonal to $|\ \rangle$.

90 Consider Robertson's inequality (4.7.15) for two Pauli operators, $A = \sigma_x$ and $B = \sigma_y$, to arrive at

$$1 - \langle\sigma_x\rangle^{\,2} - \langle\sigma_y\rangle^{\,2} - \langle\sigma_z\rangle^{\,2} \geq -\,\langle\sigma_x\rangle^{\,2}\,\langle\sigma_y\rangle^{\,2}\,.$$

Without invoking Robertson's argument, show that this is always true.

*Lorenzo Maccone (b. 1972) †Arun Kumar Pati (b. 1966)

91 Explain why

$$\big(\delta\sigma_x\big)^2 + \big(\delta\sigma_y\big)^2 \geq 1$$

is the uncertainty relation for σ_x and σ_y.

92 Next, in the spirit of Section 4.8, identify the "minimum uncertainty states" for which the equal sign holds in the inequality of Exercise 90. Do their variances saturate the uncertainty relation of Exercise 91?

93 In the Stern–Gerlach experiment in (2.1.1), assume that the silver atoms have a speed of 500 m/s and are collimated by two circular apertures of 1 mm diameter that are 50 cm apart. Estimate δX and δP and then $\delta X\,\delta P/\hbar$, with X and P referring to the transverse direction.

94 For small displacements x, the squared expectation value of the unitary displacement operator $\mathrm{e}^{\mathrm{i}xP/\hbar}$ is of the form

$$\left|\left\langle\, \mathrm{e}^{\mathrm{i}xP/\hbar}\right\rangle\right|^2 = 1 - \left(\frac{x}{L}\right)^2 + O(x^4)\,,$$

where L is the so-called *coherence length*. How is L related to the momentum spread δP? How does L compare with the position spread δX?

95 Use the three x integrals in Exercise 80 to calculate $(\delta P)^2$ for $\psi(x)$ in (4.8.10) and verify thus once more that $\delta X\,\delta P = \frac{1}{2}\hbar$ here.

96 In (4.12.7) and (4.12.8), we have the initial conditions for the time-transformation functions $\langle x,t|x',0\rangle$ and $\langle x,t|p,0\rangle$, respectively. Supplement them with the initial conditions for $\langle p,t|x,0\rangle$ and $\langle p,t|p',0\rangle$.

Chapter 5

97 For the force-free motion of Section 5.1, find the time-transformation function $\langle p,t|p',0\rangle$ from momentum states to momentum states.

98 For the force-free motion of Section 5.1, take the minimum-uncertainty wave function of (4.8.10) as the initial wave function $\psi_0(x) = \psi(x,t=0)$ and calculate $\Big\langle \big(X(0)P(0) + P(0)X(0)\big)\Big\rangle$. Then, with the previously established values for $\big\langle X(0)\big\rangle$, $\big\langle X(0)^2\big\rangle$, as well as $\big\langle P(0)\big\rangle$, $\big\langle P(0)^2\big\rangle$, determine

$\delta X(t)$. How long does it take until $\delta X(t)^2$ is twice $\delta X(0)^2$? Express this "doubling time" in terms of $\delta X(0)$ and $\delta P(0)$ without involving $\hbar$.

99 Next, express $\left\langle \frac{1}{2}\big(X(t)P(t) + P(t)X(t)\big)\right\rangle - \big\langle X(t)\big\rangle\big\langle P(t)\big\rangle$ in terms of expectation values at $t = 0$. What do you get, in particular, for the initial wave function of Exercise 98?

100 The unitary operator $U_{0,t}$ in (5.1.51) maps position bras at time $t = 0$ on those at time t. Show that the *same* $U_{0,t}$ appears in

$$\big\langle p, 0\big|U_{0,t} = \big\langle p, t\big|$$

and then conclude that $\langle \ldots, t| = \langle \ldots, 0|U_{0,t}$ for any corresponding quantum numbers.

101 As the supplement to (5.1.86), show that $\big[X, g(P)\big] = \mathrm{i}\hbar\dfrac{\partial g}{\partial P}(P)$, that is, we differentiate with respect to P by taking the commutator with X.

102 Show that $\delta(x)f(x) = \delta(x)f(0)$ in the sense of

$$\int \mathrm{d}x\ \delta(x)f(x)\,g(x) = \int \mathrm{d}x\ \delta(x)f(0)\,g(x)$$

for all functions $f(\ )$ and $g(\ )$ that are continuous at $x = 0$. Then, infer that $\delta(x-a)f(x) = \delta(x-a)f(a)$ for every permissible $f(\ )$. What is the meaning of "permissible" in this context?

103 Prove the identity for the Dirac delta function in (5.1.110) by first considering the isolated immediate vicinity of one zero of $f(x)$ by itself.

104 Invoke the completeness relation of the momentum states to verify the completeness relations (5.1.102) and (5.1.116) for the normalized energy states in (5.1.113) and (5.1.114), respectively.

105 With reference to (5.1.113) and (5.1.114), state the position wave functions $\langle x|E, \alpha\rangle$ and $\langle x|E, a\rangle$ for $\alpha = +$ or $-$ and for $a =$ even or odd and comment on them. Why is it convenient to include "i" in the definition of $|E, \mathrm{odd}\rangle$?

106 Express $\psi_E(x)$ in (5.2.14) in terms of the Airy function

$$\mathrm{Ai}(\xi) = \int_{-\infty}^{\infty} \frac{\mathrm{d}\kappa}{2\pi}\,\mathrm{e}^{-\mathrm{i}\xi\kappa}\,\mathrm{e}^{-\frac{\mathrm{i}}{3}\kappa^3}\,.$$

107 The expectation values in (5.3.93) are the diagonal elements of the matrices in (5.3.90) and (5.3.91). Determine the expectation values in (5.3.94) and (5.3.95) as the diagonal elements of the squares of these matrices.

108 Consider the three cases $a < 0 < b$, $0 < a < b$, and $a < b < 0$ to verify that

$$\int_a^b \mathrm{d}x\, \delta(x) = \frac{1}{2}\big(\mathrm{sgn}(b) - \mathrm{sgn}(a)\big)\,,$$

thereby demonstrating the assertion of (5.4.4), namely that $\frac{1}{2}\,\mathrm{sgn}(x)$ is an antiderivative of the Dirac delta function.

109 Heaviside's* unit step function $\eta(x)$ is defined by

$$\eta(x) = \left\{ \begin{array}{l} 0 \text{ if } x < 0 \\ 1 \text{ if } x > 0 \end{array} \right\} = \int_{-\infty}^{x} \mathrm{d}x'\, \delta(x')$$

and a context-dependent value at $x = 0$ which is often irrelevant. How are $\mathrm{sgn}(x)$ and $\eta(x)$ related to each other?

110 Obtain the probabilities in (5.5.50) from the probability amplitudes in (5.5.48) and (5.5.49) and confirm their unit sum.

111 Supplement the probabilities of transmission and reflection for low energies in (5.5.51) by those for high energies, $E \gg V_0$. Demonstrate that there is "above-barrier reflection," another phenomenon that contradicts classical intuition.

112 For the repulsive square-well potential in (5.5.33), find the transmission and reflection probabilities for $E = V_0$ both by solving (5.5.5) for this case and by considering the limits $V_0 > E \to V_0$ and $V_0 < E \to V_0$.

113 Supplement (5.6.5) by the equations of motion for $\sigma_x(t)$ and $\sigma_y(t)$. Then, solve them to determine the Larmor precession experienced by an atom passing through the Stern–Gerlach apparatus.

*Oliver HEAVISIDE (1850–1925)

Hints

1, 2 A combinatorial factor is needed to take into account that the order in which the clicks are observed does not matter.

3 What is XYZ?

4 What are the respective probabilities for (i) and (ii)?

5 Perhaps you want to express the columns for "$\pm$ in x" in terms of those for "$\pm$ in y."

6 The physics for the directions labeled by x, y, z is the same as that for y, z, x or z, x, y.

7 Consider the 2×2 matrix $U = \begin{pmatrix} a_1 & a_2 \\ b_1 & b_2 \end{pmatrix}$ and show first that $U^\dagger U = 1$ and then take a look at $UU^\dagger$.

8 No hint needed.

9, 10 The leading correction to the approximation in (2.6.7) sets your expectation.

11 Recognize a variation of the Zeno theme.

12 No hint needed.

13 Consider the adjoints of the statements in (2.8.13) and (2.8.14) in view of (2.8.16).

14 No hint needed.

15 You should find that $\langle a_k | b_k \rangle = 0$ and $\langle a_j | a_k \rangle = \langle b_j | b_k \rangle^*$.

16 You may want to check that $\sigma_a^\dagger = \sigma_a$ and $\sigma_a^2 = 1$ before drawing any conclusions from statements, such as $\sigma_x \sigma_y = \mathrm{i}\sigma_z$. At an intermediate step, you establish

$$\cos(\vartheta)^2 + \mathrm{i}\cos(\vartheta) = -\,\mathrm{e}^{\mathrm{i}(\varphi_x - \varphi_y)} \sin(\vartheta)^2$$

and infer that $\cos(\vartheta) = \sqrt{1/3}$, $\sin(\vartheta) = \sqrt{2/3}$.

17, 18 These are implications of (2.9.13)

19–21 Employ some of the statements in (2.9.12).

22 The particular cases of $\boldsymbol{n} = \boldsymbol{e}_x$ or $\boldsymbol{e}_y$ or $\boldsymbol{e}_z$ are available from (2.9.7) and (2.9.9).

23 Apply both sides to the eigenkets, or the eigenbras, of σ_x. Alternatively, you could first argue that $f(\sigma_x) = a + b\sigma_x$ and $f(-\sigma_x) = a - b\sigma_x$ with numerical constants a and b, and then recall the lesson of Exercise 18.

24 Express $|\uparrow_x\rangle$ and $|\downarrow_x\rangle$ in terms of $|\uparrow_z\rangle$ and $|\downarrow_z\rangle$ and then regroup the resulting four ket-bras.

25 What is $\langle a_1|A|a_2\rangle$?

26 What is the adjoint statement to $A|a\rangle = |a\rangle a$?

27 Just use the matrix that represents $\boldsymbol{e}\cdot\boldsymbol{\sigma}$ as well as the columns for the kets and the rows for the bras.

28 Combine $1 = \big(\,|\uparrow_{\boldsymbol{e}}\rangle\ |\downarrow_{\boldsymbol{e}}\rangle\,\big)\begin{pmatrix}\langle\uparrow_{\boldsymbol{e}}|\\ \langle\downarrow_{\boldsymbol{e}}|\end{pmatrix}$ with $A = 1A1$.

29 No hint needed.

30 Note the connection with Exercise 12.

31 Rely on the basic relation in (2.14.1) and the linearity of the trace.

32 Just apply the lesson of (2.14.5)–(2.14.8).

33 The trace and determinant of an operator are equal to the trace and determinant of the matrix that represent the operator with respect to an orthonormal basis of kets and bras.

34 In terms of its trace and determinant, the eigenvalues of a 2×2 matrix X are

$$\tfrac{1}{2}\mathrm{tr}(X) \pm \sqrt{\tfrac{1}{4}\mathrm{tr}(X)^2 - \det(X)}\,.$$

Establish this and then use it.

35 Begin by stating the statistical operator for this situation.

36 Note that $\mathrm{tr}\big(\rho^2\big) = \frac{1}{2}\big(\mathrm{tr}(\rho) + \boldsymbol{s}\cdot\mathrm{tr}(\boldsymbol{\sigma}\rho)\big)$.

37 You need to check only (why?) that the traces give the correct values.

38 How are the projectors $|{\uparrow_x}\rangle\langle{\uparrow_x}|$ and $|{\downarrow_x}\rangle\langle{\downarrow_x}|$ expressed in terms of σ_x?

39 A matter of checking.

40 It is enough (why?) to demonstrate the case for the $\Lambda = 1$ as well as $\Lambda = \sigma_x$, σ_y, and σ_z.

41 Just combine statements in Exercises 39 and 40 properly.

42 Follow the guidance of (2.15.4) and Section 2.16.

43 Verify that $\boldsymbol{e}$ is a unit vector and then apply a statement in Section 2.15.

44 Verify that $\sigma_z\sigma_x\sigma_z = -\sigma_x$ and $\sigma_z\sigma_y\sigma_z = -\sigma_y$ and then proceed. How is Exercise 23 helpful here?

45 Note that, for example, $|\alpha_1|^2 = |\alpha|^2 + |\gamma|^2$.

46 Recall how $|{\uparrow_x}\rangle, |{\downarrow_x}\rangle$ and $|{\uparrow_y}\rangle, |{\downarrow_y}\rangle$ are related to $|{\uparrow_z}\rangle, |{\downarrow_z}\rangle$.

47 What can you say about the expectation values of $\sigma_z^{(1)}$ and $\sigma_x^{(1)} + \mathrm{i}\sigma_y^{(1)}$?

48 Recall what you get when a Pauli operator is applied to $|{\uparrow_z}\rangle$ or $|{\downarrow_z}\rangle$.

49 It helps to identify the unitary operators that permute the Pauli operators, such as $U\sigma_x = \sigma_y U$, $U\sigma_y = \sigma_z U$, $U\sigma_z = \sigma_x U$ for a cyclic permutation.

50 No hint needed.

51 Does $A \mapsto \mathrm{tr}(A)$ map the operator A onto another operator?

52 No hint needed.

53, 54 The spectral decompositions of A and B are useful.

55 Note that $U_{ba}AU_{ab} = B$ in Exercises 53 and 54.

56 What are $f(A)^\dagger f(A)$ and $f(A)f(A)^\dagger$?

57 Recall that $f_1(A) = f_2(A)$ if $f_1(a) = f_2(a)$ for every eigenvalue a of A and and only then.

58 Write $A = A_1 + \mathrm{i}A_2$, $A^\dagger = A_1 - \mathrm{i}A_2$ with hermitian operators A_1 and A_2, the real and imaginary parts of A, and confirm that $A_1A_2 = A_2A_1$ and then proceed.

59 You should find that $\sigma_x + \mathrm{i}\sigma_y$ has only one eigenket, not two, and that $2\sigma_x + \mathrm{i}\sigma_y$ has two eigenkets that are not orthogonal.

60 Another matter of inspection.

61 The commutator in (3.5.9) is useful; you should arrive at the familiar Jacobi identity for vectors,

$$(\boldsymbol{a} \times \boldsymbol{b}) \times \boldsymbol{c} + (\boldsymbol{b} \times \boldsymbol{c}) \times \boldsymbol{a} + (\boldsymbol{c} \times \boldsymbol{a}) \times \boldsymbol{b} = 0\,.$$

62 Remember the cyclic property of the trace. For the trio of Pauli operators you should find

$$4\mathrm{i}(\boldsymbol{a} \times \boldsymbol{b}) \cdot \boldsymbol{c} = 4\mathrm{i}(\boldsymbol{b} \times \boldsymbol{c}) \cdot \boldsymbol{a} = 4\mathrm{i}(\boldsymbol{c} \times \boldsymbol{a}) \cdot \boldsymbol{b}\,.$$

63 Find $\phi(t)$ such that

$$\frac{\partial}{\partial t} + \mathrm{i}\Omega(t) = \mathrm{e}^{-\mathrm{i}\phi(t)} \frac{\partial}{\partial t} \mathrm{e}^{\mathrm{i}\phi(t)}$$

and then proceed.

64 Answer: $\omega = 1.76 \times 10^9\,\mathrm{s}^{-1} = 280\,\mathrm{MHz}$.

65 Note that $\langle\uparrow_z, t|$ and $\langle\downarrow_z, t|$ are available in (3.6.3).

66, 67 Which linear combinations of $\alpha_1(t)$, $\alpha_2(t)$, and $\alpha_3(t)$ obey simple differential equations?

68–70 No additional instructions needed.

71 Answer: $\omega_{kl} = 1.52 \times 10^{15}\,\mathrm{s}^{-1} = 2\pi \times 242\,\mathrm{THz}$.

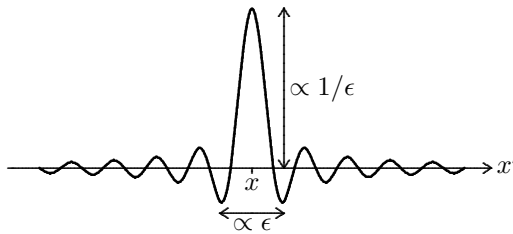
72, 73 The unitary evolution operator in (3.9.18) is useful.

74 You should find

$$\delta(x - x'; \epsilon) = \frac{1}{\epsilon\sqrt{2\pi}}\, \mathrm{e}^{-\frac{1}{2}\left(\frac{x-x'}{\epsilon}\right)^2} .$$

A gaussian function of this kind is used for the plot in (4.1.22).

75 Your graph should look like the following:



76 Exploit the simplifications offered by $\psi(-x) = \psi(x)$.

77 You expect $\langle X^n \rangle = 0$ when n is odd (why?) and can use Euler's identity in (2.5.48) for $\mathrm{e}^{\mathrm{i}kX}$ to separate the even powers from the odd powers.

78 No hint needed.

79 For example, substitute $y = 1/x$ in

$$\frac{1}{\pi} \int_{-\infty}^{\infty} \mathrm{d}y \left(\frac{\sin(y)}{y} \right)^2 = 1 .$$

Physical wave functions are not only restricted by the requirement of unit integral; we also need reasonable expectation values for important operators, such as the kinetic energy. How does this help in ensuring $\big|\psi(x)\big|^2 \to 0$ for $|x| \to \infty$?

80 Note that

$$\big\langle P^2 \big\rangle = \big\langle\ \big|P^2\big|\ \big\rangle = \int \mathrm{d}x\, \big\langle\ |x\rangle\langle x|P^2|\ \big\rangle = \int \mathrm{d}x\, \big\langle\ |P^2|x\rangle\langle x|\ \big\rangle$$

and also

$$\langle P^2 \rangle = \int \mathrm{d}x \, \langle \, |P|x\rangle\langle x|P| \, \rangle \, .$$

How do integrations by parts help in Exercise 79?

81 Different ways: With the aid of $\big|\psi(p)\big|^2$ or the lesson of Exercise 80.

82 Combine results from Exercises 77 and 81.

83 Recall the lesson of (5.1.68).

84 Establish that the variance of $\lambda^{-1}A \pm \lambda B$ is

$$\lambda^{-2}(\delta A)^2 + \lambda^2(\delta B)^2 \pm \Big(\big\langle (AB + BA) \big\rangle - 2 \, \langle A \rangle \langle B \rangle \Big)$$

and then proceed.

85 Note that $|_\perp\rangle b = (\alpha A + \beta B - a)| \, \rangle$ and $|b|^2 = b^* \langle_\perp|_\perp\rangle b = \big(|_\perp\rangle b\big)^\dagger \big(|_\perp\rangle b\big)$.

86 Realize that $\mathrm{Re}(\alpha^*\beta) = 0$ implies $\mathrm{Im}(\alpha^*\beta) = \pm|\alpha\beta|$ and make good use of the identity $|\alpha|^2(\delta A)^2 + |\beta|^2(\delta B)^2 = \big(|\alpha|\delta A - |\beta|\delta B\big)^2 + 2|\alpha\beta|\delta A \, \delta B \, .$

87 Here, realize that $\mathrm{Im}(\alpha^*\beta) = 0$ implies $\mathrm{Re}(\alpha^*\beta) = \pm|\alpha\beta|$.

88 The identity in the hint for Exercise 86 is useful again.

89 Show that $\langle_\perp{}'|(\alpha A + \beta B)| \, \rangle = \langle_\perp{}'|_\perp\rangle\langle_\perp|(\alpha A + \beta B)| \, \rangle$ with $|\langle_\perp{}'|_\perp\rangle| \leq 1$ and then proceed.

90, 91 Recall the properties of the Bloch vector $\boldsymbol{s} = \langle \boldsymbol{\sigma} \rangle$.

92 Exploit the inequality in Exercise 90 or use (4.8.1) for $C = a\sigma_x + b\sigma_y$ with $\mathrm{Re}(ab) = 0$.

93 Just argue classically to arrive at $\delta X \simeq 1\,\mathrm{mm}$ for the uncertainty in the position and $M^{-1}\delta P \simeq \dfrac{1\,\mathrm{mm}}{50\,\mathrm{cm}} \times 500 \, \dfrac{\mathrm{m}}{\mathrm{s}}$ for that in the velocity and then proceed.

94 Are you reminded of (5.1.60)?

95 Rather than evaluating the integrals, which is not difficult either, you can observe that

$$\frac{\partial}{\partial x}\psi(x) = -\frac{x}{2(\delta X)^2}\psi(x) \quad \text{and} \quad \frac{\partial^2}{\partial x^2}\psi(x) = \frac{x^2}{4(\delta X)^4}\psi(x) - \frac{1}{2(\delta X)^2}\psi(x)$$

and then conclude (how?) that

$$\langle P\rangle = \mathrm{i}\hbar\frac{\langle X\rangle}{(\delta X)^2} \quad \text{and} \quad \langle P^2\rangle = \frac{\hbar^2\langle X^2\rangle}{4(\delta X)^4} = \frac{\hbar^2}{2(\delta X)^2} - \frac{\hbar^2\langle X^2\rangle}{4(\delta X)^4}\,.$$

96 No hint needed.

97 One small step from (5.1.6).

98 You should find that $\Big\langle \big(X(0)P(0) + P(0)X(0)\big)\Big\rangle = 0$ here and that the doubling time is $M\,\delta X(0)/\delta P(0)$.

99 No hint needed.

100 Recall that $\langle p,t|x,t\rangle = \langle p,0|x,0\rangle$.

101 Adapt the argument in (5.1.87) fittingly.

102 Recall the defining property of the delta function in (4.1.14).

103 Choose a vicinity $a_j < x_j < b_j$ of x_j so small that no other zero of $f(x)$ is inside and that the linear approximation $f(x) \simeq f'(x_j)(x - x_j)$ is accurate; then show that

$$\int_{a_j}^{b_j} \mathrm{d}x\ \delta\big(f(x)\big)\,g(x) = \frac{g(x_j)}{\big|f'(x_j)\big|}$$

and proceed.

104 How are the differentials $\mathrm{d}E$ and $\mathrm{d}p$ related to each other?

105 The basic brackets in (4.6.11) are the crucial ingredients.

106 Substitute $p = (2\hbar MF)^{\frac{1}{3}}\kappa$. Verify that the product $\hbar MF$ (action $\times$ mass $\times$ force) has the correct metrical dimension of momentum-cubed.

107 You don't need to find the whole squared matrices.

108, 109 No hints needed.

110 Note the identity

$$\frac{1}{1+1/x}+\frac{1}{1+x}=1\,.$$

111 Do *not* use (5.5.50) for $E>V_0$ as these probabilities apply to $E<V_0$. Rather, find the corresponding expressions for $E>V_0$ and then consider $E\gg V_0$. At an intermediate step, the matrix

$$\begin{pmatrix}\cos\bigl(\frac{1}{2}\kappa a\bigr) & -\sin\bigl(\frac{1}{2}\kappa a\bigr)\\ \sin\bigl(\frac{1}{2}\kappa a\bigr) & \cos\bigl(\frac{1}{2}\kappa a\bigr)\end{pmatrix}$$

appears in the analog of (5.5.44), and its inverse in (5.5.45).

112 Be careful when putting $\kappa=0$ and $k=\sqrt{2MV_0}/\hbar$.

113 Consider $\sigma_x(t)\pm\mathrm{i}\sigma_y(t)$.

Index

Note: Page numbers preceded by the letters SS or PE refer to *Simple Systems* and *Perturbed Evolution*, respectively.

above-barrier reflection 208
action 127, PE33
action principle *see* quantum action principle
adiabatic approximation PE71
adiabatic population transfer PE71–73
adjoint SS10, PE1
– of a bra 34, SS10, PE1
– of a column 25
– of a ket 34, SS10, PE1
– of a ket-bra 37
– of a linear combination SS11
– of a product 37, SS18
– of an operator 37
Airy, George Bidell 155
Airy function 155, 207
algebraic completeness
– of complementary pairs PE12–14
– of position and momentum SS29–33
amplitude
– column of ∼s 92, 94
– in and out ∼s PE86
– incoming ∼ PE82
– normalized ∼ 181
– probability ∼ PE2
–– composition law for ∼∼s PE5
– probability ∼ 20–29, 33, 34, 36, 105, SS2, 13
–– time dependent ∼∼ 90–94, 102
– reflected ∼ 181, 188
– relative ∼ 181
– transmitted ∼ 181, 188
angular momentum SS93, PE25
– addition PE119–122
– and harmonic oscillators PE176
– commutation relations PE117
– eigenstates SS102
–– orthonormality SS103
– eigenvalues SS102, PE119
– intrinsic ∼ *see* spin
– ladder operators SS102, PE118
– orbital ∼ *see* orbital angular momentum
– Schwinger representation PE176
– total ∼ PE117
– vector operator SS93, PE105
angular velocity 87
– vector 91
as-if reality 55, SS27
Aspect, Alain 63
atom pairs 57–64
– entangled ∼ 57–64

– statistical operator 58, 61
axis vector SS116
azimuth SS109
– is 2π periodic SS109
azimuthal wave functions SS120
– orthonormality SS121

Baker, Henry Frederick SS36
Baker–Campbell–Hausdorff relation SS35, 36, 38
Bargmann, Valentine SS81
bases
– completeness 37, SS6, PE24
– for bras 34, 46, 59, 70, SS15, PE3, 6
– for kets 33, 46, 59, 67, 70, 99, SS15, 108, PE3, 6, 7, 68, 106
– for operators 198, 199, SS169, PE166
– for vectors 33, SS4
– harmonic oscillator ∼ 170
– mutually unbiased ∼ PE7
– orthonormal ∼ 36
– unbiased ∼ PE6, 11, 24
– unitary operator maps ∼ 75
Bell, John Stewart 4, 5
Bell correlation 5, 6, 57
Bell inequality 7, 62
– is wrong 64
– violated by quantum correlations 60, 63
Bergmann, Klaas PE73
Bessel, Friedrich Wilhelm PE107
Bessel functions, spherical ∼ *see* spherical Bessel functions
beta function integral PE37
Binnig, Gerd 191
binomial factor SS175
binomial theorem SS185
bit
– classical ∼ 68
– quantum ∼ 68
Bloch, Felix 54
Bloch vector 54
Bohr, Niels Henrik David SS113, PE14, 104, 132
Bohr energies SS125, 150
Bohr magneton 202, PE132
Bohr radius SS113
Bohr shells SS125, PE124, 160
Bohr's principle of complementarity *see* complementarity principle
Born, Max 52, 115, SS31, PE43
Born rule 51–55
Born series PE43, 48, 77, 102
– evolution operator PE43
– formal summation PE47
– scattering operator PE45
– self-repeating pattern PE45
– transition operator PE102
Born–Heisenberg commutator *see* Heisenberg commutator
Bose, Satyendranath PE142
Bose–Einstein statistics PE142
bosons PE142, 148
bound states 181
– delta potential 173–177
– hard-sphere potential PE113
– hydrogenic atoms SS116
– square-well potential 182–186
bra 31–34, SS9–15, PE1
– adjoint of a ∼ 34, PE1
– analog of row-type vector SS10
– bases for ∼s *see* bases, for bras
– column of ∼s 36, 50, 93
– eigen∼ 42–46
– infinite row for ∼ 170
– inner product of ∼s *see* inner product, of bras
– metrical dimension SS17
– orthonormal ∼s 36
– phase arbitrariness SS28
– physical ∼ SS12
– row for ∼ 96, 101
– tensor product of ∼s 59
bra-ket *see* bracket
bracket 34–38, SS12, 21
– and tensor products 59
– invariance of ∼s 75
– is inner product 35, SS11
Brillouin, Léon SS143, 158

Brillouin–Wigner perturbation theory SS143–148
Bunyakovsky, Viktor Yakovlevich SS13

Campbell, John Edward SS36
Carlini, Francesco SS158
cartesian coordinates SS105
Cauchy, Augustin-Louis SS13
Cauchy–Bunyakovsky–Schwarz inequality SS13, 167
causal link 6
causality 1, 2
– Einsteinian ∼ 6, 7
center-of-mass motion
– Hamilton operator PE140
– momentum operator PE140
– position operator PE140
centrifugal potential SS111, PE106, 110
– force-free motion SS111
classical turning point SS155, 156, 159, 165
classically allowed SS155–157, 159, 165, PE83
classically forbidden SS155–157, PE83
Clauser, John Francis 63
Clausius, Rudolf Julius Emanuel SS128
Clebsch, Rudolf Friedrich Alfred PE122
Clebsch–Gordan coefficients PE122
– recurrence relation PE135
closure relation *see* completeness relation
coherence length 206
coherent states SS81, PE62
– and Fock states SS83–84
– completeness relation SS81–83, 175, 176
– momentum wave functions SS176
– position wave functions SS77
coin tossing 68
column 25
– adjoint of a ∼ 25
– eigen∼ 96
– for ket 96, 101
– normalized ∼ 16, 25
– of bras 36, 50, 93
– of coordinates SS4
– of probability amplitudes 92, 94
– orthogonal ∼s 25
– two-component ∼s 15
column-type vector SS4
– analog of ket SS10
commutation relation
– angular momentum PE117
– ladder operators 164, SS78
– velocity PE126
commutator 83
– different times SS67, 173
– Jacobi identity 84
– position–momentum ∼ 115, SS31, 97
– product rule 84, SS31, 186
– sum rule 83, SS31
complementarity principle
– phenomenology PE27
– technical formulation PE14
completeness relation 37, 69, 74, SS6, 14, PE3, 20, 22, 24
– coherent states SS81–83, 175, 176
– eigenstates of Pauli operators 46
– Fock states SS174
– force-free states 151
– momentum states 115, SS15
– position states 106, SS15
– time dependence 81, SS38
conditional probabilities 65
constant force 153–158, PE35–36
– Hamilton operator 153, SS58, PE35
– Heisenberg equation 153, SS58, PE35
– no-force limit 155–158
– Schrödinger equation 153, 154, SS60
– spread in momentum SS59, 61
– spread in position SS59
– time transformation function SS61, 173, PE36

– uncertainty ellipse SS59
constant of motion PE130
constant restoring force SS132, 161
– ground-state energy SS134
context 67
contour integration 176, PE95
correlation
– position–momentum ∼ SS55, 59, 172
– quantum ∼s 60
Coulomb, Charles-Augustin de SS113, PE90
Coulomb potential SS113, PE90, 91
– limit of Yukawa potential PE91
Coulomb problem *see* hydrogenic atoms
cyclic permutation PE7, 16
– unitary operator PE7
cyclotron frequency PE128

de Broglie, Louis-Victor 117, SS159, PE83
de Broglie relation 117
de Broglie wavelength 117, 123, 204, SS159, PE83
deflection angle PE90, 145
degeneracy
– and symmetry 150, SS90, 116
– hydrogenic atoms SS116
– of eigenenergies 150
degree of freedom PE6
– composite ∼ PE15–16
– continuous ∼ PE17, 25
– polar angle PE25
– prime ∼ PE16
– radial motion PE25
delta function 107–109, 115, 119, 207, SS12, PE20
– antiderivative 208, SS163
– Fourier representation 119, SS17, PE21
– is a distribution 107
– model for ∼ 108, 109, SS171, PE62, 182
– more complicated argument 152, SS19
– of position operator 174
delta potential 173–181
– as a limit 186–187
– bound state 173–177
– ground-state energy 187
– negative eigenvalue 176
– reflection probability 181
– scattering states 178–181
– Schrödinger equation 175, 178
– transmission probability 181
delta symbol 69, 106, SS4, PE3, 20
– general version SS149, PE13
– modulo-N version PE180
delta-shell potential PE176
density matrix *see* statistical operator
Descartes, René 40, SS3
determinant 43, 44
– as product of eigenvalues PE68
– Slater ∼ PE144, 159
determinism 1, 2
– lack of ∼ 4, 7, 8, PE26
– no hidden ∼ 4–7
deterministic chaos 7, 8
detuning PE66
dipole moment
– electric ∼ SS154
– magnetic ∼ *see* magnetic moment
dipole–dipole interaction 98
Dirac, Paul Adrien Maurice 35, 106, SS9, PE20, 142
Dirac bracket *see* bracket
Dirac picture PE45
Dirac's delta function *see* delta function
Dirac's stroke of genius SS11
dot product *see* inner product
downhill motion 155
dyadic SS5
– matrix for ∼ SS6
– orthogonal ∼ SS9
dynamical variables 87
– time dependence PE28
Dyson, Freeman John PE47
Dyson series PE47, 48, 77

Eberly, Joseph Henry PE73
Eckart, Carl Henry PE134
effective potential energy SS111, PE158
eigenbra 42–46, 70, 81, PE2
– equation 42, 81, PE2
eigenenergies 95
eigenket 42–46, 70, 81, PE1
– equation 42, 81, PE1
– orbital angular momentum SS104
eigenvalue 42–48, 70, 81, PE2
– of a hermitian operator 77
– of a unitary operator 76
– orbital angular momentum SS104
– trace as sum of ∼s 198
eigenvector equation 42
Einstein, Albert 6, 46, PE142
Einsteinian causality 6, 7
electric field
– homogeneous ∼ SS150
– weak ∼ SS154
electron SS113
– angular momentum PE123–124
– Hamilton operator for two ∼s PE139
– in magnetic field
–– Hamilton operator PE132
electrostatic interaction SS113
energy
– and Hamilton function SS39
– Hamilton operator and ∼ values SS39
energy conservation SS156
energy eigenvalues
– continuum of ∼ 181
– discrete ∼ 181
energy spread 144, 146
entangled state 60
– maximally ∼ 200
entanglement 60
entire function SS80
equation of motion
– Hamilton's ∼ SS42, PE28
– Heisenberg's ∼ 2, 84, SS42, 43, PE27
– interaction picture PE65
– Liouville's ∼ SS44, PE29
– Newton's ∼ 2, 128
– Schrödinger's ∼ 2, 83, SS40, PE27
– von Neumann's ∼ 86, SS44
Esaki, Leo 191
Euler, Leonhard 28, 127, SS178, PE37, 152
Euler's beta function integral PE37
Euler's factorial integral SS178, PE152
Euler's identity 28
Euler–Lagrange equation 127, 130
even state (*see also* odd state) 153, 163, 178, 183, SS138
evolution operator 143, 149, SS39, PE41, 43, 74
– Born series PE43
– dynamical variables PE74
– group property PE44
– Schrödinger equation PE75
expectation value 49, SS21, 25
– of hermitian operator SS168
– probabilities as ∼s 49, SS24
expected value 49
exponential decay law PE56

face, do not lose SS134
factorial
– Euler's integral SS178, PE152
– Stirling's approximation 161
Fermi, Enrico PE48, 142, 162
Fermi's golden rule *see* golden rule
Fermi–Dirac statistics PE142
fermions PE142, 148
Feynman, Richard Phillips SS125, PE133
flipper 14, 16, 19
– anti∼ 19
Fock, Vladimir Alexandrovich 164, SS77, PE159
Fock states 164–168, SS77, 86, 89, 174
– and coherent states SS83–84
– completeness relation SS174
– generating function SS83

– momentum wave functions SS176
– orthonormality 169, SS175, 176
– position wave functions 168, 169, SS175
– two-dimensional oscillator SS94
force 148, SS130, PE85, 169
– ∼s scatter PE85
– constant ∼ *see* constant force
– Lorentz ∼ PE125
– of short range 174
– on magnetic moment 11
force-free motion 131, 135–147, 150–153, 156, PE127
– centrifugal potential SS111
– completeness of energy eigenstates 151
– Hamilton operator 131, 135, SS49, 171, PE30
– Heisenberg equation 137, 138, SS49
– orthonormal states 153
– probability flux PE173
– Schrödinger equation 135, 150, SS49, 50
– spread in momentum 138, SS54–58
– spread in position 138–141, SS53–58
– time transformation function 135–137, SS49, 50, 66, PE30
– uncertainty ellipse SS56–58
–– constant area SS58
Fourier, Jean Baptiste Joseph 119, SS3, PE21
Fourier integration SS173
Fourier transformation 119, SS2, 15, 16, 46
Fourier's theorem 119
free particle *see* force-free motion
frequency
– circular ∼ 158
– relative ∼ 48
Frullani, Giuliano PE170

g-factor PE132
– anomalous ∼ PE132
gauge function PE169
Gauss, Karl Friedrich 124, SS50, PE80
Gauss's theorem PE80
gaussian integral 125, 146, 203, SS50
gaussian moment SS3
generating function
– Fock states SS83
– Hermite polynomials 168, SS176
– Laguerre polynomials SS119
– Legendre polynomials SS122
– spherical harmonics SS124
generator 128, 130, SS92, PE32–34
– unitary transformation SS92, 129
Gerlach, Walther 9, 12, 193, PE115
Glauber, Roy Jay SS81
golden rule PE48, 51
– applied to photon emission PE53–54
– applied to scattering PE87
Gordan, Paul Albert PE122
gradient 11, SS98, PE81
– spherical coordinates SS109
Green, George SS158, PE94
Green's function PE94
– asymptotic form PE96, 97
Green's operator PE101
ground state 143, 185
– degenerate ∼s PE71
– harmonic oscillator 172, SS75
– instantaneous ∼ PE70
– square-well potential 186
– two-electron atoms PE151
ground-state energy SS131
– constant restoring force SS134
– delta potential 187
– harmonic oscillator 163
– lowest upper bound SS133
– Rayleigh–Ritz estimate SS132
– second-order correction SS141
gyromagnetic ratio *see* g-factor

half-transparent mirror 2, 4
Hamilton, William Rowan 82, SS39, 42, PE27

Hamilton function 129, SS39
– and energy SS39
Hamilton operator 82, 84, 86, 130, SS39, PE27
– and system energy SS39
– arbitrary ∼ 131
– atom–photon interaction PE53, 63
– bounded from below 143, 154
– center-of-mass motion PE140
– charge in magnetic field PE125–127
– constant force 153, SS58, PE35
– driven two-level atom PE64
– eigenbras SS86
– eigenstates 149
– eigenvalues 149
–– degeneracy 150, PE128
– electron in magnetic field PE132
– equivalent ∼s 85–86
– force-free motion 131, 135, SS49, 171, PE30
– harmonic oscillator 158, 163, 164, SS66, 68, 72–74, 85
– hydrogenic atoms SS113
– matrix representation PE68
– metrical dimension SS39
– photon PE52, 63
– relative motion PE140
– rotation PE116
– spherical symmetry SS108
– three-level atom PE72
– time-dependent ∼ 84, SS40, PE28
– time-dependent force SS62
– two electrons PE139
– two-dimensional oscillator SS89, 106
– two-electron atoms PE149
– two-level atom PE52, 68
– typical form 132, 148, SS41, 155, PE80
–– virial theorem SS128
Hamilton's equation of motion SS42, 43, PE28
hard-sphere potential PE110, 174
– bound states PE113
– impenetrable sphere PE113
– low-energy scattering PE112
harmonic oscillator 158–172, SS66
– anharmonic perturbation SS141
– bases 170
– eigenenergies 162
– energy scale SS75
– ground state 172, SS75
– ground-state energy 163
– ground-state wave function 164
– Hamilton operator 158, 163, 164, SS66, 68, 72–74, 85
–– eigenkets SS76
–– eigenvalues SS76
– Heisenberg equation 158, SS66, 71, 84
– ladder operators 166, SS76, 174, PE52, 176
– length scale 159, SS73, 75
– momentum scale SS73, 75
– momentum spread 172, SS174
– no-force limit SS70, 71
– position spread 172, SS174
– Schrödinger equation 158, SS68
– time transformation function SS67, 70, 84, 85, 88, 174
– two-dimensional isotropic ∼ *see* two-dimensional oscillator
– virial theorem SS129
– wave functions 163, 165, 167–169
–– orthonormality 169
– WKB approximation SS160, 180
Hartree, Douglas Rayner PE159
Hartree–Fock equations PE159
Hausdorff, Felix SS36
Heaviside, Oliver 208, SS163
Heaviside's step function 208, SS163
Heisenberg, Werner 2, 84, 115, 123, SS31, 42, 172, PE19, 27
Heisenberg commutator 115, SS31, 38, PE19, 22
– for vector components SS97
– invariance 115
Heisenberg equation (*see also* Heisenberg's equation of motion) 84, SS42, 43, 74, PE27

– and Schrödinger equation SS171
– constant force 153, SS58, PE35
– force-free motion 137, 138, SS49
– formal solution 149
– general force 148
– harmonic oscillator 158, SS66, 71, 84
– solving the ∼s 149
– special cases SS44
– Stern–Gerlach apparatus 192
– time-dependent force SS62
Heisenberg picture PE45
Heisenberg's
– equation of motion (*see also* Heisenberg equation) 2, SS42, 43, PE27
– formulation of quantum mechanics 115
– uncertainty relation 123
Heisenberg–Born commutator *see* Heisenberg commutator
helium SS154, PE148
helium ion SS113
Hellmann, Hans SS125, PE133
Hellmann–Feynman theorem SS125–127, 140, 149, 179, PE133
Hermite, Charles 77, SS10, 176, PE4
Hermite polynomials 162, 168–170
– differential equation 162
– generating function 168, SS176
– highest power 162
– orthogonality 169
– Rodrigues formula 168
– symmetry 162
hermitian conjugate *see* adjoint
hermitian operator 76–77, SS168, PE4
– eigenvalues 77, PE165
– reality property SS168, PE165
Hilbert, David 78, SS11, 182, PE2
Hilbert space 78, SS11, PE2
Hilbert–Schmidt inner product 78, SS182
Hund, Friedrich Hermann 191
hydrogen atom (*see also* hydrogenic atoms) SS113
hydrogen ion PE148, 154
hydrogenic atoms
– and two-dimensional oscillator SS115
– axis vector SS116
– Bohr energies SS125, 150
– Bohr shells SS125
– bound states SS116
– degeneracy SS116
– eigenstates SS115
– eigenvalues SS115
– Hamilton operator SS113
– Laplace–Runge–Lenz vector SS117
– mean distance SS127
– natural scales SS113
– radial wave functions SS121
–– orthonormality SS122
– scattering states SS116
– Schrödinger equation SS113
– total angular momentum PE123
– virial theorem SS129
– wave functions SS121–124
– WKB approximation SS180

identity operator 37, SS13, PE3
– infinite matrix for ∼ 171
– square of ∼ 106
– square root of ∼ PE65, 67
– trace 53, 61
indeterminism *see* determinism, lack of ∼
indistinguishable particles PE139
– kets and bras PE141
– permutation invariance of observables PE139
– scattering of ∼ PE144–148
infinitesimal
– change of dynamics PE30, 32
– changes of kets and bras PE32
– endpoint variations 128
– path variations 130
– rotation 87, SS99
– time intervals 80

– time step SS39
– transformation SS92
– unitary transformation 81
– variation SS69
– variations of an operator PE36
inner product 35, SS5, 11, PE2
– Hilbert–Schmidt ∼ 78, SS182
– of bras 35, 77, SS12
– of columns 35, SS5
– of kets 35, 77, SS11
– of operators 78, SS182
– of rows 35, SS5
integrating factor 154, SS60, 157
interaction picture PE45, 65
interference 156, PE5
inverse, unique ∼ 19

Jacobi, Carl Gustav Jacob 84, SS100
Jacobi identity SS100
– for commutators 84
– for vectors 212
Jeffreys, Harold SS158

Kelvin, Lord ∼ *see* Thomson, William
Kennard, Earle Hesse 123, SS172
Kepler, Johannes SS116
Kepler ellipse SS116
ket 31–34, SS9–15, PE1
– adjoint of a ∼ 34, PE1
– analog of column-type vector SS10
– bases for ∼s *see* bases, for kets
– basis ∼s 35, 70
– column for ∼ 96, 101
– eigen∼ 42–46
– entangled ∼ 60
– infinite column for ∼ 170
– inner product of ∼s *see* inner product, of kets
– metrical dimension SS17
– normalization of ∼s 35, 69
– orthogonality of ∼s 35, 69
– orthonormal ∼s 36
– phase arbitrariness SS28
– physical ∼ SS10, 12
– reference ∼ 33
– row of ∼s 36, 50
– tensor product of ∼s 59, PE144
ket-bra (*see also* operator) 37, 38, SS21
– adjoint of a ∼ 37
– tensor products of ∼s 59
kinetic energy 131, SS41, 105, 111
kinetic momentum 130
Kirkwood, John Gamble SS169
Kirkwood operators SS169
Kramers, Hendrik Anthony SS158
Kronecker, Leopold 69, SS4, PE3
Kronecker's delta symbol *see* delta symbol

ladder operators 166
– angular momentum SS102, PE118
– commutation relation 164, SS78
– differentiation with respect to ∼ SS78
– eigenbras SS79
– eigenkets SS77
– eigenvalues SS77
– harmonic oscillator 166, SS76, 174, PE52, 176
– lowering operator 166
– normal ordering SS80, 174, 175
– orbital angular momentum SS102, 104
– raising operator 166
– two-dimensional oscillator SS93, 94, PE129
Lagrange, Joseph Louis de 127, PE157
Lagrange function 128
Lagrange parameter PE157
Lagrange's variational principle 127
Laguerre, Edmond SS108
Laguerre polynomials SS119
– generating function SS119
– Rodrigues formula SS119
Lamb, Willis Eugene PE58
Lamb shift PE58

Langer, Rudolph Ernest SS162
Langer's replacement SS163
Laplace, Pierre Simon de SS117, PE60
Laplace differential operator SS178
Laplace transform PE60
– inverse ∼ PE61
– of convolution integral PE61
Laplace–Runge–Lenz vector SS117
Laplacian *see* Laplace differential operator
Larmor, Joseph 90
Larmor precession 90, 208
Legendre, Adrien Marie SS122, PE106
Legendre function SS122
Legendre polynomials SS122, PE106
– generating function SS122
– orthonormality PE107
Lenz, Wilhelm SS117
light quanta 2
Liouville, Joseph SS44, PE29
Liouville's equation of motion SS44
Lippmann, Bernard Abram PE46
Lippmann–Schwinger equation PE77, 98, 101
– asymptotic form PE98
– Born approximation PE100
– exact solution PE104
– iteration PE46
– scattering operator PE46
lithium ion SS113, PE148
locality 6, 7
Lord Kelvin *see* Thomson, William
Lord Nelson *see* Rutherford, Ernest
Lord Rayleigh *see* Strutt, John William
Lorentz, Hendrik Antoon PE60, 125
Lorentz force PE125
Lorentz profile PE60

Maccone, Lorenzo 205
Maccone–Pati uncertainty relation 205
magnetic field 19
– charge in ∼
–– circular motion PE130–132
–– Hamilton operator PE125–127
–– Lorentz force PE127
–– probability current PE177
–– velocity operator PE125
– homogeneous ∼ 14, 27, 38, 39, 79, 90, PE127
– inhomogeneous ∼ 9
– potential energy of magnetic moment in ∼ 10
– vector potential PE125
magnetic interaction energy 84, 98, 202, PE132
magnetic moment 10, 90, PE115
– force on ∼ 11
– potential energy of ∼ in magnetic field 10
– rotating ∼ PE115
– torque on ∼ 10
many-electron atoms
– binding energy PE161
– outermost electrons PE164
– size PE163
matrices
– 2×2 ∼ 15
– infinite ∼ 170–171
– Pauli ∼ *see* Pauli matrices
– projection ∼ 26
– square ∼ 15
– transformation ∼ SS7
Maxwell, James Clerk 1
Maxwell's equations 2
mean value 49, 111, SS19
measurement
– disturbs the system 31
– equivalent ∼s 80
– nonselective ∼ 55–57
– result 48, PE2
– with many outcomes 68–73
Mermin, Nathaniel David 66
Mermin's table 66
mesa function SS24
metrical dimension
– bra and ket SS17, PE22
– Hamilton operator SS39

– Planck's constant SS18, 39
– wave function SS17
model 191
momentum 129
– canonical ∼ PE126
– classical position-dependent ∼ SS156
– kinetic ∼ 130, PE126
momentum operator 112–114, 130, SS21
– differentiation with respect to ∼ 207, SS31, 78, 98, 170, PE22
– expectation value 119–120, SS55
– functions of ∼ 117, SS21
– infinite matrix for ∼ 171
– spread 138, 172, SS54
– vector operator SS97
momentum state SS15
– completeness of ∼s 115, SS15
– orthonormality of ∼s 115, SS15
motion
– to the left 150, 180, SS156, PE84
– to the right 150, 180, SS156, PE84

Nelson, Lord ∼ *see* Rutherford, Ernest
Newton, Isaac 1, 158, PE91
Newton's equation of motion 2, 128, 158
Newton's theorem PE179
Noether, Emmy SS92
Noether's theorem SS92
normal ordering SS80, 174, 175
– binomial factor SS175
normalization 27
– force-free states 152
– state density PE52
– statistical operator SS27
– wave function 110, SS2, 3, 12

observables PE1
– complementary ∼ PE5, 11, 14, 26
– mutually exclusive ∼ PE5
– undetermined ∼ PE27
odd state (*see also* even state) 153, 163, 178, 183, SS138
operator (*see also* ket-bra) 37, PE1
– adjoint of an ∼ 37
– antinormal ordering SS81
– bases for ∼s *see* bases, for operators
– characteristic function SS170
– equal ∼ functions 72
– evolution ∼ *see* evolution operator
– function of an ∼ 71, 72, PE4
–– unitary transformation of ∼∼ 201, SS34, PE18
–– varying a ∼∼ PE36, 37, 169
– hermitian ∼ *see* hermitian operator
– identity ∼ *see* identity operator
– infinite matrix for ∼ 170
– inner product of ∼s *see* inner product, of operators
– logarithm of an ∼ PE170
– normal ∼ 200, 201, SS182
– normal ordering SS80
– not normal ∼ 201
– ordered ∼ function SS30–33, PE12–14, 166
– Pauli ∼s *see* Pauli operators
– Pauli vector ∼ *see* Pauli vector operator
– projection ∼ *see* projector
– reaction ∼ *see* reaction operator
– scalar ∼ SS102
– scattering ∼ *see* scattering operator
– spectral decomposition 72, 81, SS21, PE3, 4
– spread 120
– statistical ∼ *see* statistical operator
– unitary ∼ *see* unitary operator
– variance 121
– vector ∼ SS102, 177
optical theorem PE102, 175

orbital angular momentum SS95, 98, PE116, 117
– commutators SS99–102
– eigenkets SS104
– eigenstates SS102
– eigenvalues SS95, 102, 104
– ladder operators SS102, 104
– vector operator SS98
–– cartesian components SS93, 98, 99
ordered exponential SS32
orthogonality 35, PE3
– of kets 35, 69
orthohelium 69
orthonormality 69, 81, SS6, PE3, 20, 22, 24
– angular-momentum states SS103
– azimuthal wave functions SS121
– Fock states 169, SS175, 176
– force-free states 153
– Legendre polynomials PE107
– momentum states 115, SS15
– position states 106, SS12, 15
– radial wave functions SS121, 122
– spherical harmonics SS122
– time dependence SS38
overidealization 109, SS10

partial waves
– for incoming plane wave PE107
– for scattering amplitude PE109
– for total wave PE108
particle
– identical ∼s *see* indistinguishable particles
– indistinguishable ∼s *see* indistinguishable particles
Pati, Arun Kumar 205
Pauli, Wolfgang 40, SS105, PE115, 142
Pauli matrices 38–40, PE177
Pauli operators 38–40, 47
– functions of ∼ 40–41
– nonstandard matrices for ∼ 197, 198
– standard matrices for ∼ 40, 198
– trace of ∼ 53
– uncertainty relation 206
Pauli vector operator 40, 52, SS105, PE115
– algebraic properties 40
– commutator of components 88
– component of ∼ 44
Peierls, Rudolf Ernst PE104
permutation
– cyclic ∼ PE7, 16
persistence probability PE49
perturbation theory 149, PE37
– Brillouin–Wigner *see* Brillouin–Wigner perturbation theory
– for degenerate states SS148–155
– Rayleigh–Schrödinger *see* Rayleigh–Schrödinger perturbation theory
phase factor 28, SS61
phase shift 178
phase space SS56
phase-space function SS30, 33
phase-space integral SS32, 33, 164
photoelectric effect 46
photon 2
– Hamilton operator PE52, 63
photon emission PE52–62
– golden rule PE53–54
– probability of no ∼ PE59
– Weisskopf–Wigner method PE54–60
photon mode PE54, 55
photon-pair source 4
Placzek, George PE104
Planck, Max Karl Ernst Ludwig 82, SS3, PE19
Planck's constant 82, SS3, PE19
– metrical dimension SS18, 39
Poisson, Siméon Denis PE23
Poisson identity PE23
polar angle SS109
polar coordinates SS105
polarizability SS155
position operator 110–112, SS20
– delta function of ∼ 174

– differentiation with respect to ∼ 148, SS31, 78, 98, 170, PE22
– expectation value 111, SS55
– functions of ∼ 111, SS20
– infinite matrix for ∼ 171
– spread 138, 172, SS54
– vector operator SS97
position state SS15
– completeness of ∼s 106, SS15
– orthonormality of ∼s 106, SS12, 15
position–momentum correlation SS55, 59, 172
potential energy SS41, 109
– effective ∼ SS111, PE158
– localized ∼ PE82, 86
– separable ∼ PE87, 173
potential well 174
power series method 159–162
prediction *see* statistical prediction
principal quantum number SS115
principal value PE58
– model for ∼ PE62
probabilistic laws 4
probabilistic prediction *see* statistical prediction
probability 20–29, SS1, 13, PE2, 79
– amplitude 20–29, 33, 34, 36, 105, SS2, 13, PE2
–– column of ∼s 92, 94
–– time dependent ∼∼ 90–94, 102
– as expectation value 49, SS24
– conditional ∼ 65, SS2
– continuity equation PE80, 172
– current density PE79, 81, 99
–– charge in magnetic field PE177
– density 110, SS1, PE79
– flux of ∼ PE79
– for reflection 181, 188, 190
– for transmission 181, 188, 190
– fundamental symmetry 73, SS13, PE2
– local conservation law PE80
– of no change 141–148
–– long times 142
–– short times 142, 143
– transition ∼ *see* transition probability
probability operator *see* statistical operator
product rule
– adjoint SS18
– commutator 84, SS31, 186
– transposition SS5
– variation PE181
projection 26
– matrices 26
– operator *see* projector
projector 26, 41, 47, 110, 111, SS182, PE8, 10
– on atomic state PE52
– to an x state 174
property, objective ∼ 13, 60

quantum action principle PE32–39
quantum state estimation 65
qubit 68

Rabi, Isidor Isaac PE53
Rabi frequency PE53, 54, 66
– modified ∼ PE67, 78
– time dependent ∼ PE63, 71
radial density SS122
radial quantum number SS107
radial Schrödinger equation SS111, PE106
Rayleigh, Lord ∼ *see* Strutt, John William
Rayleigh–Ritz method SS131–138, PE151, 157
– best scale SS134
– excited states SS137–138
– scale-invariant version SS136
– trial wave function SS132
Rayleigh–Schrödinger perturbation theory SS138–143, 146, 148
reaction operator PE170
reflection
– above barrier ∼ 208
– symmetry 150
relative frequency 48
relative motion

– Hamilton operator PE140
– momentum operator PE140
– position operator PE140
residue 177, PE96
Riemann, Georg Friedrich Bernhard PE21
Ritz, Walther SS132
Robertson, Howard Percy 122
Robertson's uncertainty relation 122, 204
Rodrigues, Benjamin Olinde 168, SS119
Rodrigues formula
– Hermite polynomials 168
– Laguerre polynomials SS119
Rohrer, Heinrich 191
rotation SS8, 90, PE25
– consecutive ∼s SS100–101
– Hamilton operator PE116
– internal ∼ PE116
– orbital ∼ PE116
– rigid ∼ PE116
– unitary operator SS93, 99
row 25
– eigen∼ 96
– for bra 96, 101
– of coordinates SS4
– of kets 36, 50
row-type vector SS4
– analog of bra SS10
Runge, Carl David Tolmé SS117
Rutherford, Ernest (Lord Nelson) PE91
Rutherford cross section PE91
Ry *see* Rydberg constant
Rydberg, Janne SS114, PE150
Rydberg constant SS114, PE150

s-wave scattering PE110–114
– delta-shell potential PE176
– hard-sphere potential PE112
scalar product (*see also* inner product) SS5
scaling transformation SS130
– and virial theorem SS130
scattering 181, PE82–114
– Born series PE102
– by localized potential PE86
– cross section
–– Coulomb potential PE91
–– golden-rule approximation PE89
–– Rutherford ∼∼ PE91
–– separable potential PE173
–– Yukawa potential PE91
– deflection angle PE90, 106, 175
– elastic ∼ PE88
– elastic potential ∼ PE89
– electron-electron ∼ PE144–147
– forward ∼ PE102
– golden rule PE87
– in and out states PE92
– incoming flux PE88
– inelastic ∼ PE88
– interaction region PE86, 97
– low-energy ∼ *see* *s*-wave scattering
– of *s*-waves *see* *s*-wave scattering
– of indistinguishable particles PE144–148
– right-angle ∼ PE146
– separable potential PE104
– spherically symmetry potential PE89
– transition matrix element PE89
– transition operator PE100–175
–– separable potential PE105
scattering amplitude PE98
– and scattering cross section PE100
– and scattering phases PE109
– and transition operator PE101
– partial waves PE109
scattering cross section PE88
– and scattering amplitude PE100
– and scattering phases PE110
scattering matrix PE86
scattering operator PE44, 76
– Born series PE45
– equation of motion PE76
– integral equation PE45
– Lippmann–Schwinger equation PE46

scattering phase PE109
scattering states 181
– delta potential 178–181
– hydrogenic atoms SS116
– square-well potential 187–191
Schmidt, Erhard 78, SS182
Schrödinger, Erwin 2, 60, 83, 114, 117, 159, SS1, 40, 138, 172, PE25, 27
Schrödinger equation (*see also* Schrödinger's equation of motion) 2, 83, SS40, 43, PE27
– and Heisenberg equation SS171
– driven two-level atom PE64, 65
– evolution operator PE75
– for bras 83, SS40, PE27
– for column of amplitudes 94
– for kets 83, SS40, PE27
– for momentum wave function 131, 133
– for position wave function 131, 132, SS41
– force-free motion 135, SS49, 50
– formal solution SS86
– harmonic oscillator SS68
– initial condition 131
– radial ∼ *see* radial Schrödinger equation
– solving the ∼ SS45
– time independent ∼ 96, 100, SS156, PE83
–– constant force 153, 154
–– delta potential 175, 178
–– force-free motion 150
–– harmonic oscillator 158
–– hydrogenic atoms SS113
–– spherical coordinates SS110
–– square-well potential 187
–– two-dimensional oscillator SS106, 115
– time transformation function SS45, 46
– two-level atom PE55
Schrödinger picture PE45
Schrödinger's
– equation of motion (*see also* Schrödinger equation) 83, SS40, PE27
– formulation of quantum mechanics 114
– uncertainty relation 204, SS172
Schur, Issai PE167
Schur's lemma PE167
Schwarz, Hermann SS13
Schwinger, Julian SS81, PE15, 33, 46, 162, 177
Schwinger representation PE176
Schwinger's quantum action principle PE32–39
Scott, John Moffett Cuthbert PE162
Scott correction PE162
Segal, Irving Ezra SS81
selection 13
– successive ∼s 13, 17
selector 13, 16
selfadjoint *see* hermitian
SG *acronym* Stern–Gerlach
short-range force 174
sign function 173, SS169
– and step function SS182
silver atom 9, 10, 48, 57, 87, 97, 202, 206, PE115
single-photon counter 3
singlet PE123, 143, 146
– projector on ∼ PE179
Slater, John Clarke PE144
Slater determinant PE144, 159
solid angle PE87
solid harmonics SS123
spectral decomposition 72, 81, SS21, PE3, 13, 21
speed PE126
speed of light PE125
spherical Bessel functions PE107
– asymptotic form PE107
spherical coordinates 44, SS108, 109, PE89, 94
– gradient SS109
– local unit vectors SS109
– position vector SS109

– Schrödinger equation SS110
spherical harmonics SS122, PE106
– differential equation SS123
– examples SS123
– generating function SS124
– orthonormality SS122
– symmetry SS123
spherical wave
– incoming ∼ PE108, 109
– interferes with plane wave PE102
– outgoing ∼ PE98, 99, 108
spin 10, SS105, PE117
spin–orbit coupling PE137
spin–statistics theorem PE142
spread 120
– and variance 121
– geometrical significance 145
– in energy 144, 146
– in momentum 138, SS54
– in position 138, SS54
– vanishing ∼ 145
spreading of the wave function 139, 141, SS54, 56, PE183
square-well potential 182–191
– above-barrier reflection 208
– attractive ∼ 187
– bound states 182–186
– count of bound states SS180
– ground state 186
– reflection probability 188, 190
– repulsive ∼ 187
– scattering states 187–191
– Schrödinger equation 187
– transmission probability 188, 190
– tunneling 187–191
Stark, Johannes SS153, PE137
Stark effect
– linear ∼ SS153
– quadratic ∼ SS154
state (*see also* statistical operator) 52
– bound ∼s *see* bound states
– coherent ∼s *see* coherent states
– entangled ∼ *see* entangled state
– even ∼ *see* even state
– Fock ∼s *see* Fock states
– mixed ∼ SS27, 29
– odd ∼ *see* odd state
– of minimum uncertainty 123–126
– of the system 1, 2, SS3
– pure ∼ SS28, 29
– reduction 64–65
–– is a mental process 65
– scattering ∼s *see* scattering states
– vector 33
state density PE51
– normalization PE52
state of affairs 1, SS3, 9, 23
state operator *see* statistical operator
stationary phase 156
statistical operator 121, SS23, 26, PE26
– blend 55, SS27
–– as-if reality 55, SS27
– Born rule 51–55
– for atom pairs 58, 61
– inferred from data SS26
– mixture 55, SS27
–– many blends for one ∼ 55, SS27
– nature of the ∼ 65
– normalization SS27
– positivity SS27
– represents information 56, 65, SS26
– time dependence 86, SS44, PE28
– time-dependent force SS173
statistical prediction 4, 8, 65, SS2, 26
– verification of a ∼ 22, SS2
step function 208, SS163
– and sign function SS182
Stern, Otto 9, 12, 193, PE115
Stern–Gerlach
– apparatus 47, 48, 69, 195, 196
–– entangles 193
–– equations of motion 192
–– generalization 69
– experiment 9–12, 27, 46, 191–193, 206
–– displacement 193
–– Larmor precession 208

–– momentum transfer 193
– magnet 10, 12, 196, PE115
– measurement 21, 23, 27, 32, 147
– successive ∼ measurements 12–15
Stirling, James 161
Stirling's approximation 161
Strutt, John William (Lord Rayleigh) SS124, 132, 138
surface element PE79
swindle PE50
symmetry
– and degeneracy 150, SS90, 116
– reflection ∼ 150

Taylor, Brook 112
Taylor expansion 156
Taylor's theorem 112
tensor product 59–60
– and brackets 59
– of bras 59
– of identities 61
– of ket-bras 59
– of kets 59, PE119, 144
Thomas, Llewellyn Hilleth PE162
Thomas–Fermi energy PE162
Thomson, William (Lord Kelvin) SS122
three-level atom PE71
– Hamilton operator PE72
time dependence
– dynamical ∼ 84, SS43, 44, 88, PE28
– parametric ∼ 84, SS40, 43, 44, 88, PE28
time ordering PE46
– exponential function PE47
time transformation function 133–134, SS45–47, 87, PE29
– as a Fourier sum SS88
– constant force SS61, 173, PE36
– dependence on labels SS51
– force-free motion 135–137, SS49, 50, 66, PE30
– harmonic oscillator SS67, 70, 84, 85, 88, 174
– initial condition 134, 206, SS45, 46
– Schrödinger equation SS45, 46
– time-dependent force SS63–66
– turning one into another 134, SS46
time-dependent force
– Hamilton operator SS62
– Heisenberg equation SS62
– spread in momentum SS63
– spread in position SS63
– statistical operator SS173
– time transformation function SS63–66
Tom and Jerry 57–65
torque on magnetic moment 10
trace 49–51, SS22, PE14
– as diagonal sum 50
– as sum of eigenvalues 198, PE68
– cyclic property 51, SS25
– linearity 50, SS23
– of ordered operator SS32
– of Pauli operators 53
– of the identity operator 53, 61
transformation function 116, 118, SS16, PE22
– time ∼ *see* time transformation function
transition PE47, 48
– frequency 97, 103, PE49
– operator (*see also* scattering, transition operator) PE52, 100
– probability PE48, 49, 51
– rate PE49, 51, 54, 87
translation
– unitary operator SS93
transposition SS4, 10
– of a product SS5
trial wave function SS132
triplet PE123, 143, 146
– projector on ∼ PE179
tunnel diode 191
tunnel effect 191
tunnel transistor 191
tunneling microscope 191
tunneling probability 191

two-dimensional oscillator SS89–95, PE129
– and hydrogenic atoms SS115
– degeneracy SS90
– eigenstates SS90, 93–95, 106
– Fock states SS94
– Hamilton operator SS89, 106
– ladder operators SS93, 94, PE129
– radial wave functions SS121
–– orthonormality SS121
– rotational symmetry SS90
– Schrödinger equation SS106, 115
– wave functions SS117–121
two-electron atoms PE148–158
– binding energy PE154
– direct energy PE155
– exchange energy PE155
– ground state PE151
– Hamilton operator PE149
– interaction energy PE152
– single-particle energy PE152
two-level atom
– adiabatic evolution PE68–71
– driven ∼ PE62–71
–– Hamilton operator PE64
–– Schrödinger equation PE64, 65
– frequency shift PE58, 62
– Hamilton operator PE52, 68
– instantaneous eigenstate PE71
– periodic drive PE66
– projector on atomic state PE52
– resonant drive PE65
– Schrödinger equation PE55
– transition operator PE52
– transition rate PE54, 58, 62

unbiased bases PE6
uncertainty ellipse SS56–58
– area SS57, 173
uncertainty principle 123
uncertainty relation 120–123, 204, 205
– for Pauli operators 206
– Heisenberg's ∼ 123
–– and Kennard 123
–– and Schrödinger SS172
–– more stringent form SS172
– of the Maccone–Pati kind 205
– physical content 123
– Robertson's ∼ 122, 204
– Schrödinger's ∼ 204
uncertainty, state of minimum ∼ 123–126, 206
unit
– atomic-scale ∼s 82
– macroscopic ∼s 82
unit matrix 26
unit vector 44
unitary operator 73–77, SS19, 168, PE4
– eigenvalues 76, PE165
– for cyclic permutations PE7
– for shifts PE24
– maps bases 75
– period PE7
– rotation SS93, 99
– transforms operator function 201, SS34, PE18
– translation SS93
unitary transformation SS38, 92
– generator SS92, 129
uphill motion 155

variance (*see also* spread) 121, 138, 145
– geometrical significance 145
vector 32
– coordinates of a ∼ 32
– state ∼ 33
vector potential
– asymmetric choice PE128
– symmetric choice PE128
vector space SS9
velocity 138, PE80, 125
– commutation relation PE126
virial theorem SS128, 130
– and scaling transformation SS130
– harmonic oscillator SS129
– hydrogenic atoms SS129
von Neumann, John 86, SS44, PE29
von Neumann equation 86, PE29

wave function 105, SS1
– its sole role SS2
– momentum ∼ 118, SS2
–– metrical dimension SS17
– normalization 110, SS2, 3, 12
– position ∼ 118, SS2
–– metrical dimension SS17
– spreading *see* spreading of the wave function
– trial ∼ SS132
wave train 123
wave–particle duality 46
Weisskopf, Victor Frederick PE54
Wentzel, Gregor SS158
Wentzel–Kramers–Brillouin *see* WKB
Weyl, Claus Hugo Hermann SS34, PE12, 15, 166
Weyl commutator SS34, PE12
Weyl's operator basis PE166, 181
Wigner, Eugene Paul SS143, PE54, 134
Wigner–Eckart theorem PE134
WKB *acronym* Wentzel–Kramers–Brillouin
WKB approximation SS155–165
– harmonic oscillator SS160, 180
– hydrogenic atoms SS180
– reliability criterion SS159
WKB quantization rule SS160, 161, 165
– in three dimensions SS162
– Langer's replacement SS163

Yukawa, Hideki PE90
Yukawa potential PE90
– double ∼ PE174
– scattering cross section PE91

Zeeman, Pieter PE137
Zeeman effect PE137
Zeilinger, Anton 63
Zeno effect 29–31, 148
Zeno of Elea 31